

# NASA Contractor Report 172118 - vol-1

National Aeronautics and Space Administration (NASA)/  
American Society for Engineering Education (ASEE)  
Summer Faculty Fellowship Program--1988

## Volume 1

Richard B. Bannerot, Editor  
*University of Houston - University Park*  
*Houston, Texas*

Stanley H. Goldstein, Editor  
*University Programs Office*  
*Lyndon B. Johnson Space Center*  
*Houston, Texas*

(NASA-CR-172118-Vol-1) NATIONAL AERONAUTICS  
AND SPACE ADMINISTRATION (NASA)/AMERICAN  
SOCIETY FOR ENGINEERING EDUCATION (ASEE)  
SUMMER FACULTY FELLOWSHIP PROGRAM 1988,  
VOLUME 1 (NASA) 199 F

N89-20058  
--THRU--  
N89-20072  
Unclass  
C19C025

CSCI 05I G3/99

Grant NGT-44-005-803

February 1989



National Aeronautics and  
Space Administration

Lyndon B. Johnson Space Center  
Houston, Texas

## **PREFACE**

The 1988 Johnson Space Center (JSC) National Aeronautics and Space Administration (NASA) / American Society for Engineering Education (ASEE) Summer Faculty Fellowship Program was conducted by the University of Houston and JSC. The 10-week program was operated under the auspices of the ASEE. The program at JSC, as well as the programs at other NASA Centers, was funded by the Office of University Affairs, NASA Headquarters, Washington, D.C. The objectives of the program, which began in 1965 at JSC and in 1964 nationally, are

1. To further the professional knowledge of qualified engineering and science faculty members
2. To stimulate an exchange of ideas between participants and NASA
3. To enrich and refresh the research and teaching activities of participant's institutions
4. To contribute to the research objectives of the NASA Centers

Each faculty fellow spent at least 10 weeks at JSC engaged in a research project commensurate with his/her interests and background and worked in collaboration with a NASA / JSC colleague. This document is a compilation of the final reports on the research projects done by the faculty fellows during the summer of 1988. Volume 1 contains reports 1 through 14, and volume 2 contains reports 15 through 26.

# CONTENTS

## Volume 1

1.	Ahmed, Selina: "Comparison of Soviet and U.S. Space Food & Nutrition Programs" .....	1-1
2.	Allen, Robert: "Production System Chunking in SOAR: Case Studies in Automated Learning" .....	2-1
3.	Bartram, Peter N.: "Development of a Software Interface for Optical Disk Archival Storage for a New Life Sciences Flight Experiments Computer" .....	3-1
4.	Bell, Christopher and Lachman, Roy: "Computer Technologies and Institutional Memory" .....	4-1
5.	Berry, Frederick C.: "Development of Parallel Algorithms for Electrical Power Management in Space Applications" .....	5-1
6.	Botros, Nazeih M.: "Automatic Voice Recognition Using Traditional and Artificial Neural Network Approaches" .....	6-1
7.	Casserly, Dennis M.: "Development of an Atmospheric Monitoring Plan for Space Station" .....	7-1
8.	Chang, Ing: "Model Formulation of Non-Equilibrium Gas Radiation for Hypersonic Flight Vehicles" .....	8-1
9.	Davis, John E.: "A Model for Plasma Volume Changes During Short Duration Spaceflight" .....	9-1
10.	de Korvin, Andre: "An Evidential Approach to Problem Solving When a Large Number of Knowledge Systems is Available" .....	10-1
11.	Elam, Reid P.: "Effects on Motor Unit Potentiation and Ground Reaction Force From Treadmill Exercise" .....	11-1
12.	Flynn, George J.: "An Assessment of the Micrometeoritic Component in the Martian Soil" .....	12-1
13.	Geer, Richard D.: "Evaluation of Available Analytical Techniques for Monitoring the Quality of Space Station Potable Water" .....	13-1
14.	Kelly, Frederick A.: "Model Evaluation, Recommendation, and Prioritizing of Future Work for the Manipulator Emulator Testbed" .....	14-1

## Volume 2

15.	Knopp, Jerome: "Optical Joint Correlation Using the Deformable Mirror Device" .....	15-1
16.	McLauchlan, Robert A.: "Intelligent Control of Robotic Arm / Hand Systems for the NASA EVA Retriever Using Neural Networks" .....	16-1

17.	Nechay, Bohdan R.: "Maladjustment of Kidneys to Microgravity: Design of Measures to Reduce the Loss of Calcium" .....	17-1
18.	Nelson, Dennis O.: "Evaluation of Crustal Recycling During the Evolution of Archean-Age Matachewan Basaltic Magmas" .....	18-1
19.	Nelson, Kerri L.: "Lithospheric Buckling and Intra-Arc Stresses: A Mechanism for Arc Segmentation" .....	19-1
20.	Nerheim, Rosalee: "Toward an Image Compression Algorithm for the High-Resolution Electronic Still Camera" .....	20-1
21.	Odell, Patrick L.: "Equating an Expert System to a Classifier in Order to Evaluate the Expert System" .....	21-1
22.	Randhawa, Manjit S.: "Feasibility of Using High Temperature Superconducting Magnets and Conventional Magnetic Loop Antennas to Attract or Repel Objects at the Space Station" .....	22-1
23.	Starks, Scott A.: "Visual Perception and Grasping for the Extravehicular Activity Robot" .....	23-1
24.	Tan, Arjun: "Estimation of Velocity Perturbations in Satellite Fragmentation Events" .....	24-1
25.	Tolliver, C.L.: "Time-Domain Imaging" .....	25-1
26.	Wiley, Ronald J.: "Arc Jet Diagnostics Tests" .....	26-1

**PRECEDING PAGE BLANK NOT FILMED**

**N89 - 20059**

**COMPARISON OF SOVIET AND U.S. SPACE FOOD  
AND NUTRITION PROGRAMS**

**Final Report**

**NASA/ASEE Summer Faculty Fellowship Program--1988**

**Johnson Space Center**

<b>Prepared by:</b>	<b>Selina Ahmed, Ph.D.</b>
<b>Academic Rank:</b>	<b>Associate Professor</b>
<b>University &amp; Department:</b>	<b>Texas Southern University Department of Home Economics Houston, TX 77004</b>
<b>NASA/JSC</b>	
<b>Directorate:</b>	<b>Space &amp; Life Sciences</b>
<b>Division:</b>	<b>Man-Systems Division</b>
<b>Branch:</b>	<b>Flight Systems Branch</b>
<b>JSC Colleague:</b>	<b>Charles T. Bourland, Ph.D.</b>
<b>Date:</b>	<b>August 22, 1988</b>
<b>Contract Number:</b>	<b>NGT44-001-800</b>

## **ABSTRACT**

This report compares the Soviet Space Food and Nutrition programs with those of the U.S. The Soviets established the first Space Food programs in 1961, when one of the Soviet Cosmonauts experienced eating in zero gravity.

A Soviet scientist recently developed "trophology" - the study of living systems which includes assimilation of nutrients. Trophology is expected to permit nutritional advances beyond the "Balanced Diet" concept of satisfying ongoing metabolic needs. The concept expands and enriches the old concepts with the new findings in space nutrition. The Soviet scientists have conducted a number of studies regarding the concepts of nutrition assimilation and increased nutritional needs in long duration missions.

U.S. Space Food and Nutrition programs have been developed over the past twenty-five years. From the early days of Mercury and Gemini to future Space Station requirements, the U.S. Space Food and Nutrition programs have progressively improved.

This study indicates that some major differences exist between the two space food and nutrition programs regarding dietary habits. The major differences are in recommended nutrient intake and dietary patterns between the cosmonauts and astronauts. The intake of protein, carbohydrates and fats are significantly higher in cosmonaut diets compared to astronauts. Certain mineral elements such as phosphorus, sodium and iron are also significantly higher in the cosmonauts' diets. Cosmonauts also experience intake of certain unconventional food and plant extract to resist stress and to increase stamina.

## **COMPARISON OF SOVIET AND U.S. SPACE FOOD AND NUTRITION PROGRAMS**

### **INTRODUCTION**

The successful conquest of space was facilitated considerably by the creation of a life support system for cosmonauts and astronauts. One of the more important elements of life support is nourishment for the spacecraft crews. Collective efforts of technologists, nutritionists, engineers, physicians and microbiologists have developed a solution for these problems. Foods not only provide astronauts or cosmonauts nutrition for normal vital activities, but also provide psychological fulfillment.

For more than thirty years, Russian food industry technologists, scientists and specialists have studied and developed various forms, types and groups of food products. This resulted in a specialized food production industry to develop food for the cosmonauts. Nine food products with very limited storage lives were developed and tested for Yuriy Gagarin in 1961. One of the first tasks of the first man in space was to attempt eating in zero gravity. A food ration was then proposed for flights lasting from one to four days. The products were canned meats, the first dinner foods, and fruit juices in aluminum tubes, and bite-sized pieces of bread. All were considered to have very limited shelf lives. (20, 21)

The primary purpose of the U.S. Space Food and Nutrition Program is to provide nutritionally balanced diets to the astronauts during space flights and to gather physiological information and conduct nutritional studies designed to assess the effects of space flights on nutrient metabolism and its effects on crew performance. Over the past twenty-five years, the designs of food systems for spacecraft have been developed from the experience of past food systems. The foods used on U.S. space flights have been comprised of a wide variety of foods which have been specially processed and/or packaged to adapt them to zero gravity consumption. Mercury food was experimental and transient. Aluminum tubes were used for the semi-solid foods. The first manned flight of the Gemini program, Gemini 3, lasted less than five hours, but four experimental meals were aboard to test. The meals consisted of dehydrated foods in cubes, dry mix and freeze-dried products providing up to 2900 Kcal/crewmember. The Shuttle food system has over 150 items for astronauts to choose from when selecting a menu. The foods are classified by the method of preservation, namely, rehydratable, thermostabilized, natural form, and intermediate moisture. (19)

### **OBJECTIVES**

1. To study the historical background and progression of Soviet Space Food and Nutrition programs.
2. To acquire knowledge about cosmonauts' daily dietary intake and nutritional supplements during longer duration missions.
3. To study their physiological information and nutritional research and assess the effects of space flight on nutrient metabolism on crew performance.
4. To compare the Soviet Food and Nutrition program to the U.S.

## REVIEW OF THE LITERATURE

In the Soviet space programs, nutrition is given an important place in the life support systems for spacecraft crews. The "Soyuz-9" spacecraft, which was launched on July 1, 1970, with a crew of Andriyan Nikolayev and Vitaliy Sevastyanov, completed an 18-day flight. Their mission carried a device for heating food products in aluminum tubes. This made it possible to prepare hot foods onboard the spacecraft. In 1974, aboard the "Salyut-4" space station, a device was used to simultaneously heat products in aluminum tubes, and bread, which significantly improved the food intake of the spacecraft crews. (1, 20)

In the Salyut-6 missions, the food technologists, scientists, and other specialists expanded the variety of food products, increased their quality, and extended their shelf lives. A six-day ration was developed which provided a daily four-meal course (first and second breakfast, lunch and dinner). The Salyut-6 mission ration in 1977 had over 70 different products including twenty types of meat, fish, poultry, and cheese in 100gm cans, together with pureed first and second courses (10 types) in aluminum tubes. There were fourteen types of dehydrated (freeze dried) first and second courses and garnishes for the canned meats, which could be rapidly reconstituted by adding hot water, Bakery products, and desserts made from twelve types of semi-dehydrated fruits were also available. The drink selection included fruit juices, tea, coffee, milk, and fermented milk products.

In Soviet Space Food and Nutrition programs, the make-up of the diets are developed on the basis of the flight duration, the complexity of the program, the anticipated energy use, the storage, heating and dispensing equipment, the water reclamation systems, and other equipment. In long term space flights, the relative monotony concerning diet does become annoying to some degree. In order to prevent the monotony of the diets and uplift the psychological aspects of life, some specially prepared fruits and vegetables were transported on an experimental basis. The shipment of fresh apples, onions, and garlic also was transported at the request of cosmonauts G. Crechko and Yu Romenenko during their Soyuz mission. The use of fresh fruits and vegetables was expanded. Oranges, lemons, melons, honey, cranberries, and even fresh cherries, were sent at the request of V. Ryumin during his Soyuz mission. The cargo spacecraft were used to send fresh fruits and vegetables to Vladimir Lyakhov and Aleksander Aleksandrov, who stayed three months in space during Soyuz T-9 mission in 1983.

The food service provisions aboard the Salyut-6 mission consisted of the following elements: Container for stowing and storing foods, tables to make meals, an electric warmer, place settings, a device to recycle, measure, and dispense hot and cold water, and containers for disposal of packages and left over foods. (15, 9)

**TABLE I - MENU EXAMPLE (COSMONAUTS) (4)**

**MENU EXAMPLE (COSMONAUTS)**

List of foods for a one-day menu on Salyut-6:

First Breakfast:

Chicken with Prunes	100g
Bread	45g
Candy	50g
Coffee with Milk	150g

Lunch:

Cottage Cheese with Pureed Black Currants	165g
Honeycake	45g
Black Currant Juice with Pulp	50g

Dinner:

Sauerkraut Soup	165g
Roast Beef with Mashed Potatoes	57.5g
Bread	45g
Prunes with Nuts	60g
Candied Fruit	50g
Coffee with Sugar	24g

Supper:

Chicken in Tomato Sauce	165g
Bread	45g
Cheese	100g
Tea with Sugar	23g

V. P. Bychkov, et. al, 1986, researched the diets of three Salyut-7 prime crews. The Salyut-7 cosmonauts experienced the best nutritional systems of the Soviet Space Nutrition programs. The rations of the Salyut-7 crews had a caloric value of 3150 Kcalories and were well balanced, containing all the nutritional requirements. The food system included cosmonauts rations, containers for serving and storing food, a dining table, an electric food heater, utensils and dishes, devices for regenerating water and measuring hot and cold water to packages of freeze-dried food, containers for disposing of waste, and a refrigerator for storing fruit and vegetables. The food had equivalent food value of diets of Salyut-6, but consisted primarily (65%) of freeze-dried products, rehydrated with hot or cold water. (5, 17, 20)

In a study by V. P. Bychkov, et. al., 1981, concerning the adequacy of the protein supplied in the diet of crews of Salyut-6, it was found that the physical status of the cosmonauts and the parameters of nitrogen metabolism after the missions were indicative of adequate protein intake. Examination of nitrogen metabolism in the crew of the first main expedition aboard Salyut-6 (96-day mission) revealed that, in spite of individual fluctuations in the excretion of the end products of nitrogen metabolism in the post-flight

period, their range was within the pre-flight levels. Urinary excretion of the total nitrogen was 11.1 gm per day post-flight and 11.6 gm per day pre-flight. On the 3rd and 4th post-flight days, it dropped to 9.0 gm per day, then reverted to the base level on the 5th and 7th days. (8)

In the second expedition aboard Salyut-6 (140-day mission), the post-flight excretion of the end product of nitrogen metabolism was close to pre-flight levels. On the 3rd post-flight day, there was a 40% decrease in excretion of total nitrogen, 21% decrease for urea, and 44% decrease for ammonia, with retention of normal proportion. There was a tendency toward normalization of these changes on subsequent post-flight days. (8)

I. G. Popov, et. al, 1982, conducted a study to determine the effect of a 48-day flight on the blood amino acids content in the crew of Salyut-5. It was found that changes in essential and non-essential amino acids ratios, e.g., methionine and cysteine, are of major concern. They observed that the changes in amino acids metabolism could be due to several nutritional factors. The transition from the pre-flight diet to the inflight diet could lead to a reduction in the essential amino acids content. A decreased intake of dairy products, eggs, leguminous products, freshly prepared food, and an excess intake of foods which had gone through severe heat processing and drying. They suggested that this offsets the amount of amino acids in foods and their accessibility to digestion. They also suggested that changes in living conditions may change the synthesis of amino acids. Similar studies were conducted by I. G. Popov, et. al., 1983, to measure the amino acids levels in blood of cosmonauts during the 185-day Salyut-6 flight. It was found that appreciable decreases in the concentration of most amino acids occurred in both cosmonauts. The threonine and cysteine levels in blood plasma of both cosmonauts were below the bottom of the normal range cited in the Russian Medical Encyclopedia. Alanine and histidine for the commander, and methionine, isoleucine, and arginine for the flight engineer, were also below the normal range. The demonstrated changes in plasma amino acids were attributed to similar factors cited by Popov in 1982. Recommended countermeasures are also similar, such as organizing nutrition rehabilitation for cosmonauts by increasing all amino acid intake, particularly methionine, cysteine, valine, histidine, tyrosine, arginine, aspartic, and glutamic acids. (23, 24)

K. V. Smirnov, et. al., 1982, studied the state of the digestive system following long-term space flights. Those studies included: gastric, pancreatic and intestinal enzymes. These studies demonstrated consistent changes in the digestive system function. The depth and severity were related to duration and the condition of the space flights. It was found that weightlessness plays an effective role in these changes. There was a noticeable change of increased activity of gastric pepsinogen and pancreatic lipase after the first day of the 96-day mission. They were normalized by the 25th day of the readaptation period for the 96- and 140-day flights, and at the 43rd day after the 175-day space flights. They suggested inflight use of preventive measures and conditioning of cosmonauts to develop changes in enzyme secretion in the gastrointestinal tract during the readaptation period. The marked changes were not noted on the first day after the 185-day space flight, in comparison with the 175-day flight, due to the preventive measures of adequate nutrition used during the readaptation period. (27)

V. P. Bychkov, et. al., 1982, studied the effects of hypokinesia on man's nutritional status. They found that during and after the bed rest studies adequate nutrition and certain unidentified nutrients can serve as efficient countermeasures against metabolic changes, such as weight loss, negative nitrogen, phosphorus, potassium, and sodium balance. The nutrient rehabilitation led to a positive nitrogen balance, increase in the utilization of vitamin C, B1, B6, and faster termination of compensatory nitrogen retention, as compared to subjects in the control group. Physical exercise with adequate nutrition was found to be the most effective means of preventing changes observed under hypokinetic conditions. (3)

I.A. Radayeva, et. al., 1982, conducted research regarding the biological value and shelf life of cultured dairy products in the diet of cosmonauts. Freeze-dried cultured milk products, yogurt with sugar, and yogurt with fruits and berries were examined for shelf life. It was determined that the realistic shelf life was 12 months at  $20^{\circ} \pm 5^{\circ} \text{C}$  and 18 months at  $1-4^{\circ} \text{C}$ . They considered these foods to have a high biologic value with respect to their protein. The leucine/isolucine ratio is 2:1 for acidophilus paste, 2:3 for sweetened yogurt, 2:4 for fruit and berry yogurt. It was reported that freeze-dried cultured dairy products may be quite beneficial in the diet of cosmonauts and were recommended to be incorporated into space diets. (26)

N. G. Bogdanov, et. al., 1986, studied vitamin levels in cosmonauts during pre-flight training and after completion of short-term space flights. They reported a statistically significant decrease in the excretion of a number of vitamins during the post-flight period. They suggested that it was due to an increased vitamin metabolism which leads to an increased need for them during post-flight. (2)

## DISCUSSION

Soviet specialists believe that both a regular meal schedule and carefully selected diet are important for the maintenance of overall conditioning in space. Nutrition is thought to have a synergistic effect with other countermeasures on the control of adaptive changes, such as musculoskeletal strength and mass losses. It has been observed that cosmonauts can actually gain weight in space if exercise is combined with vitamins, calcium supplements, appealing foods, and appetite stimulators, such as onions, garlic and spices. (4, 16)

## USSR SPACE FOOD AND NUTRITION CONCEPTS

### FOOD

In the 30th minute of the 1961 mission, Yuri Gagarin ate and drank. This became the first evidence of the possibility of eating, chewing, and swallowing liquid and solid food in weightlessness. Daily ration of the cosmonauts on Vostok and Vostok-2 contained about 2800 Kcal, including 100 gm of protein, 118 gm of fat and 308 gm of carbohydrates. Foods which were used in the flights were packaged in dispenser tubes and included soups, cottage cheese, as well as drinks: coffee, cocoa, juices. The rations had limited shelf lives without refrigeration (up to 5 to 6 days). Meat products in packets had to be prepared directly prior to flight. (20, 21)

In the Soyuz missions, foods with long shelf life, such as bread, were baked in the form of small "one-bite" rolls to prevent crumbs. Meat products such as ham, steak, and veal, were also included in the diet. The daily dietary intake for cosmonauts in Soyuz had a 3-day menu cycle with four meals per day. The Soyuz mission also included dehydrated boiled meat, and Soyuz-9 cosmonauts were the first to heat meals at 60° to 70° C and drink from the dispenser tubes. (20)

Cosmonauts P. Popvich and Yu Artyukhin on Salyut-3 were the first to test dehydrated products, rehydrated with recovered water. These experiments were continued in Salyut-4. Dehydrated products now comprise up to 20% of the food rations. The second crew of the Salyut-4 took additional food products of limited shelf life, as well as bread, coffee, and tea for the first time in the transport ship. The unmanned "Progress" transport ship is in widespread use to deliver fresh food rations to the orbital station. (20)

The water supply system on board the spacecraft functions in conjunction with the food systems. The supply of drinking water in the orbital station is produced at a rate of up to two liters per man per 24 hours. (17)

The most recent and satisfactory cosmonaut nutritional system on Soviet manned space flights was the one developed for Salyut-7. There is a buffet table on Salyut for eating food, as well as a set of table accessories, a food heater, facilities for sanitary cleaning of the table accessories, and bags for leftovers and packaging. The calorie content of the daily diet was increased up to 3200 Kcal to combat the negative consequences of weightlessness and physical training exercises during the missions. (5, 16)

## NUTRITION

-Dr. Oleg Gazenko, 1987, in an address to the U. N. committee on the peaceful uses of outer space, discussed the concepts regarding digestive physiology. He described "a recently developed branch of medicine - trophology" as the study of general principles of fundamental vital processes of living systems, such as ingestion, processing, and assimilation of nutrients. Several discoveries in this area were outlined: New information regarding the immunological defense of the small intestine, importance of dietary fiber in digestive functioning and overall health, and microbiology of the digestive tract in medical support of long-term space flight. (12)

Dr. Gazenko suggests that trophology will advance the nutritionists' knowledge beyond the accepted concept of "Balanced Diet", which means satisfying on-going metabolic needs, to the new concepts of "Adequate Nutrition." The new concept does not really replace the previous concepts but expands and enriches them with new findings in the field of space nutrition. (12)

In research findings, decreases observed in essential amino acids in cosmonauts after the 211-day mission led researchers to the conclusion that the pre-flight diet should be supplemented with methionine and aspartic acid and inflight and post-flight diets with seven essential amino acids plus cysteine, arginine, proline, and aspartic acid. (6, 11)

Soviet scientists believe that increased exercise regimens require an additional intake of calories to maintain proper energy balance. They also believe that supplements consisting of vitamins, amino acids and minerals promote the retention of fluids and electrolytes. They have shown that pro-

per dietary combinations can also help to regulate the digestive and enzymatic changes associated with stresses of space flights. (22, 25)

The Soviet scientists also believe that the plant extract (Eleutherococcus), exercise and high calcium diet are three major ways by which reduction in calcium loss is possible in the spacecraft crews. They also use the plant extract to resist stress and to increase stamina among the cosmonauts. (10)

### PACKAGING AND PRESERVATION

Snack products, bread products and appetizers, as well as sweet pastry products and fruits, were packaged for the Soviet space missions in film packets made of viscotene, a clear plastic film material similar to polyethylene. Some of them were vacuum packed. The ration of cosmonauts in the Soyuz ship included new shelf stable foods with a long shelf life, and pureed and liquid products in dispenser tubes. Meat products - ham, steak, and veal - were prepared in the form of meals preserved in metal cans. The sweet products included chocolate candy, prunes with nuts, and honey ginger bread, all in film packages. Some of the products, in the form of briquettes, were covered in edible film. The dehydrated boiled meat included in the menu was vacuum packed in a film. (9, 17, 20, 25)

### U.S. SPACE FOOD AND NUTRITION CONCEPTS

The food systems planned for the U.S. Space Station are detailed in reference (28).

### SUMMARY AND CONCLUSION

The research study shows that the successful conquest of space, either for astronauts or cosmonauts, depends on the collective efforts of technologists, engineers, physicians, microbiologists, nutritionists and psychologists.

According to the Congressional Research Service Report prepared by Hon. Ernest F. Hollings, May 1988, the Soviets have continued to make steady strides toward their goals of having a permanently occupied space station in Earth orbit. The Soviets hold a commanding lead in the operational use of crews. They have introduced two new launch vehicles and continue to develop a space shuttle and space plane. (13, 14)

A comparison time line of U.S. and Soviet space missions is shown in the Appendix.

Cosmonaut Romanenko, after his 326-day flight in space in 1987, stated that Mars is getting nearer and nearer to Earth. The Soviets also conducted a variety of studies dealing with new concepts regarding nutritional assimilation and increased needs in long duration missions. However, when comparing the research data, these reports are equivocal regarding countermeasures taken, the increased needs, and types of tests used in the analyses. (14)

The present research study indicates that food for U.S. space flight has improved steadily throughout the space programs. From the early days of Mercury and Gemini to future Space Station requirements, the U.S. Space Food and Nutrition program has progressively improved. The U.S. Space Food and Nutrition programs are more advanced in terms of their food packaging,

preservation techniques and flexible menu patterns for long duration missions. (28) However, some major differences exist between the two space nutrition programs regarding the respective dietary habits. These differences are shown in Tables 2 and 3.

TABLE 2 - DIFFERENCES IN RECOMMENDED NUTRIENT INTAKE OF COSMONAUTS AND ASTRONAUTS

<u>NUTRIENTS</u>	<u>COSMONAUTS</u>	<u>ASTRONAUTS</u>
Kilocalories	3200 Kcal	2300-3100 Kcal
Protein	1.5g/kg Bwt. (140g)	8g/kg Bwt.
Fat	1.4g/kg Bwt. (100g)	1.3g/kg Bwt. (93g)
Carbohydrates	4.5g/kg Bwt. (395g)	4.8g/kg Bwt. (350g)
Phosphorus	1.7g	.8g
Sodium	4.5g	3.5g
Iron	50mg	18mg

TABLE 3 - OTHER DIFFERENCES IN FOOD INTAKE OF COSMONAUTS AND ASTRONAUTS

<u>FOOD</u>	<u>COSMONAUTS</u>	<u>ASTRONAUTS</u>
Plant Extract (Eleutherococcus) 500mg/day or 1.00g every other day	Used to increase stamina to resist stress	None
-Garlic	As food seasoning	None
Vodka	Small amount	None
Brandy	Small amount	None
Fresh Fruits and Vegetables	Supplied by Progress ship	Stored 16 hours before launch
Onion, Dill, Parsley	Cultivated in on-board vegetable garden (18)	None
Multivitamin	Supplement--twice/day	Optional (Shuttle)
Undevit	Vitamin--twice/day	None
Aerovit	Vitamin--twice/day	None
Essential Amino Acids (Methionine)	Supplements/increased amount	None
Glutamic Acid Decamerit	Supplement	None

The cosmonauts' nutrient intake is probably higher than the astronauts'. The mean daily inflight nutrient consumption per person during Shuttle STS-1 through STS-61C indicate higher intake of nutrients by the astronauts compared to the RDA (Appendix). The difficulty exists in estimating the nutrient intake since Shuttle crews are not required to maintain a food intake log. (29)

#### FUTURE RECOMMENDATIONS FOR U.S. SPACE NUTRITION PROGRAMS

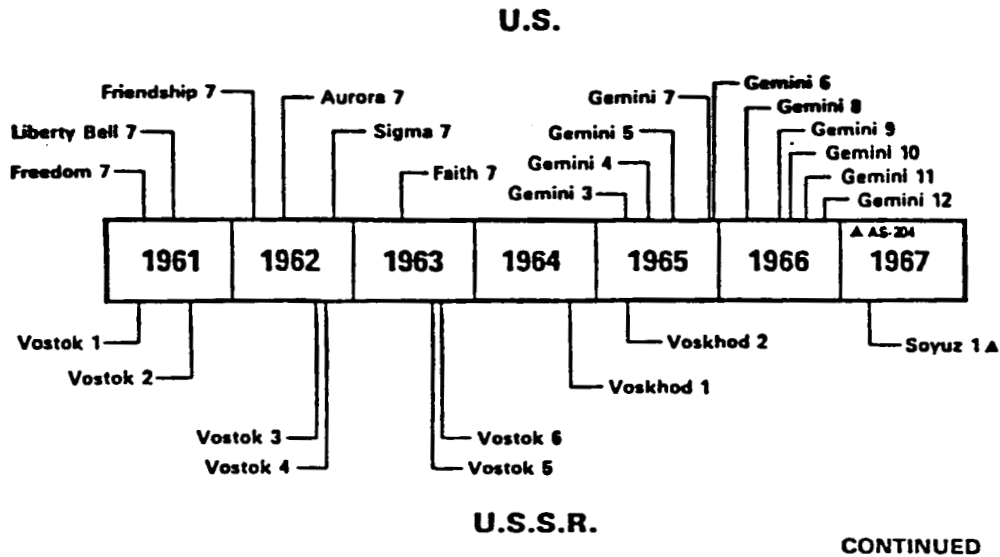
1. Determine nutrient requirements for longer duration missions.
2. Understand the increased essential amino acids requirements during pre-flight, inflight, and post-flight diets.
3. Determine the effects of plant extract as a stimulant.
4. Determine the beneficial effects of alcohol for longer duration missions in space.
5. Study the regulatory effects of proper dietary combinations on digestive and enzymatic changes which are associated with stresses of long duration flights.
6. Study the accepted concept of "Balanced Diet", which means satisfying only ongoing metabolic needs, to the new concept of "Adequate Nutrition", which does not really replace the previous concepts, but expands and enriches them with new findings in the field of space nutrition.

## REFERENCES

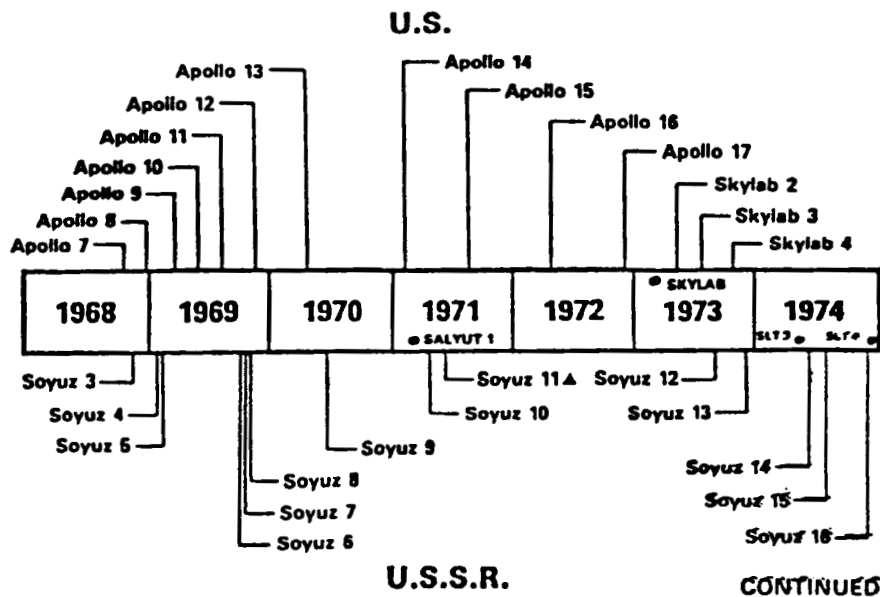
1. Bluth, B. J., et. al., Soviet Space Station as Analogs, Second Edition. California State University, Northridge, May 1987.
2. Bogdanov, N. G., Gvozdoval, L. G., Belakovskiy, M. S., Smirnova, A. N., Blazheyevich, N. V., Tarabaniko, V. M., Yuzhanskaya, M. G., Pastushkova, L. K. H., Zaburkina, T. G., and Pereverzeva, O. G., Vitamins Levels in Cosmonauts During Pre-Flight Training and Completion of Short-Term Space Flights. USSR Space Life Science Digest, Issue No. 8, 1986.
3. Bychkov, V. P., Gryaznova, V. N., Kalandarov, S., Kasatkina, A. G., Korshunova, V. A., Markaryan, M. V., Sivuk, A. K., and Khokhlova, O. S., Effects of Hypokinesia on Man's Nutritional Status. USSR Reports: Space Biology and Aerospace Medicine, Vol. 16, No. 2, pp 34-39, 1982.
4. Bychkov, V. P., Kalandarov, S., Markaryan, M. V., Radchanko, N. D., Stepchikov, K. A., and Frumkin, M. L., Diet of Crews of Three Main Expeditions Aboard Salyut-6 Orbital Station. USSR Report No. 5, pp 21-25, November 1982.
5. Bychkov, V. P., Kalandarov, S., Kochetkova, A. N., Sedova, Y. E. A., Ushakov, A. S., and Frumkin, M. L., Diet of Cosmonauts of the Three Salyut-7 Prime Crews. Space Biology and Aerospace Medicine Digest, Issue 14, pp 220-221, 1986.
6. Bychkov, V. P., Vlasova, T. F., Gryaznova, V. N., Sedova, Y. E. A., Sivuk, A. K., Tretyakova, V. A., and Ushkov, A. S., Biological Value of Protein in Food Allowance of Salyut Orbital Station Crews. USSR Report: Space Biology and Aerospace Medicine (JPRS-USB-86-006), 1986.
7. Bychkov, V. P., Ushakov, A. S., Kalandarov, S., Markaryan, M. V., Sedova, Y. E. A., Sivuk, A. K., and Khokhlova, O. S., Diet of Crews in Salyut-6 Orbital Station. USSR Report: Space Biology and Aerospace Medicine, Vol. 16, No. 2, pp 11-15, 1982.
8. Bychkov, V. P., Sivuk, A. K., and Borodulina, I. I., Conformity with Human Requirements of Protein Contained in the Rations for Crews of the Salyut-6 Orbital Station. USSR Report: Space Biology and Aerospace Medicine, Vol. 15, No. 1, 1981.
9. Frumkin, M. L., Kudrova, R. V., Kuznetsova, L. I., Akin'shina, G. G., Yershova, A. A., and Gurova, L. A., Special Features of Technology for Producing Food Products for Cosmonauts. USSR Reports: Space, JPRS, Arlington, VA.
10. Fulder, S., The Drug That Builds Russians. New Scientist, pp 576, August 21, 1980.
11. Gazenko, O. G., and Yegorov, A. D., 175-Day Space Flight; Some Results of Medical Research. Vestnik Akademii Nauk SSSR No. 9, pp 49-58, September 1980.
12. Gazenko, O. G., Space Medicine - New Approaches in the Theory and Practice of General Medicine. Presentation of the Soviet Delegation at the 24th Session of the Science and Technology Subcommittee of the U.N. Committee of Peaceful Uses of Outer Space, United Nations, New York, February 1987.
13. Haeseler, D., Maintaining a Space Station Spaceflight. Vol. 28, pp 426-429, December 1986.

14. Hollings, E. F., Soviet Space Programs: 1981-87. Piloted Space Activities, Launch Vehicles, Launch Sites, and Tracking Support: Committee on Commerce, Science and Transportation, United States Senate, Part I: May 1988.
15. Kidger, N., Salyut-6 Mission Report - Part 3. Space Flight, Vol. 22, pp 146-154, April 4, 1980.
16. Kidger, N., Salyut Mission Report. Space Flight, Vol. 25, pp 122-125, March 3, 1983.
17. Konovalov, B., New Food Selection System for Cosmonauts in Salyut-7. USSR Report No. 19, pp 15-16, January 1983.
18. Konovalov, B., New Plant-Growing Experiments on 'Salyut-7'. USSR Report: Space, No. 19, pp 15-16 (JPRS-82771).
19. Klicka, M. V., and Smith, M. C., Food for U.S. Manned Space Flight. Technical Report, Natick/TR81/019 Food Engineering Laboratory, NASA, 1982.
20. Maltsev, A., and Frumkin, M., Food Specialists Discuss Cosmonauts' Diet. USSR Report: Space (JPRS-USP-84-003).
21. Perov, V., and Skolenko, Y., Space Home, Scientific Horizon "From Breakfast to Dinner". Nauka i Religiiia No. 4, pp 2-6, 1982, in Russian. (Translated by Khendker K. Rob, Senior Engineer, Lockheed Engineering and Science Company, JSC, June 1988)
22. Popov, I. G., and Latskevich, A. A., Some Distinctions Referable to Amino Acid Levels in Blood of Cosmonauts Who Participated in 185-Day Flight. JPRS, Arlington, VA, 1983.
23. Popov, I. G., and Latskevich, A. A., Effect of 48-Day Flight on Blood Amino Acid Content in the Crew of Salyut-5. USSR Report: Space Biology and Aerospace Medicine, Vol. 16, No. 2, 1982.
24. Popov, I. G., and Latskevich, A. A., Free Amino Acids in Blood of Salyut-5 Crew Before and After 21-Day Mission (Second Expedition). USSR Reports: Space Biology and Aerospace Medicine, Vol. 17, No. 1 (JPRS-83007), 1983.
25. Popov, I. G., Food and Nutrition for Long Missions in Space. USSR Report, No. 13, pp 40-44, October 28, 1981.
26. Radayeva, I. A., Rossikhina, G. A., Usacheva, V. A., Poyarkova, G. S., and Shul'kina, S. P., Biological Value and Shelf Life of Cultured Dairy Products in the Diet of Cosmonauts. USSR Reports: Space Biology and Aerospace Medicine, Vol. 16, No. 2, pp 29-33, 1982.
27. Smirnov, K. V., Syrykh, G. D., Gegen'kov, V. I., Goland-Ruvina, L. G., Medkova, I. L., and Voronin, L. I., State of Digestive System Following Long-Term Space Flights. USSR Reports: Space Biology and Aerospace Medicine, Vol. 16, No. 2, pp 23-28, 1982.
28. Space Station Food Supply and Service System (FSSS) Technical Report, Vol. 1, LEMSCO - 24210, September 1987.
29. Stadler, C. R., Rapp, R. M., Bourland, C. T., and Fohey, M. F., Space Shuttle Food System Summary. Technical Report, NASA/Johnson Space Center, Houston, TX, in press.

## TIME LINE OF U.S. AND U.S.S.R. PILOTED SPACE MISSIONS

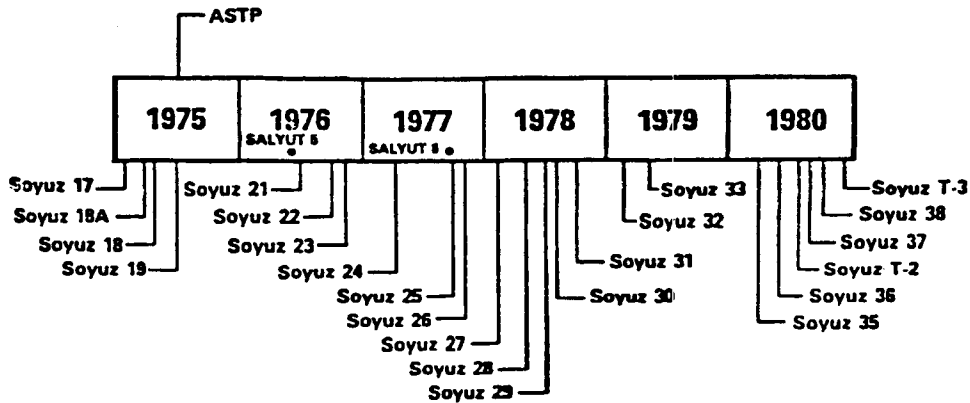


## TIME LINE OF U.S. AND U.S.S.R. PILOTED SPACE MISSIONS (continued)



## TIME LINE OF U.S. AND U.S.S.R. PILOTED SPACE MISSIONS (continued)

### U.S.

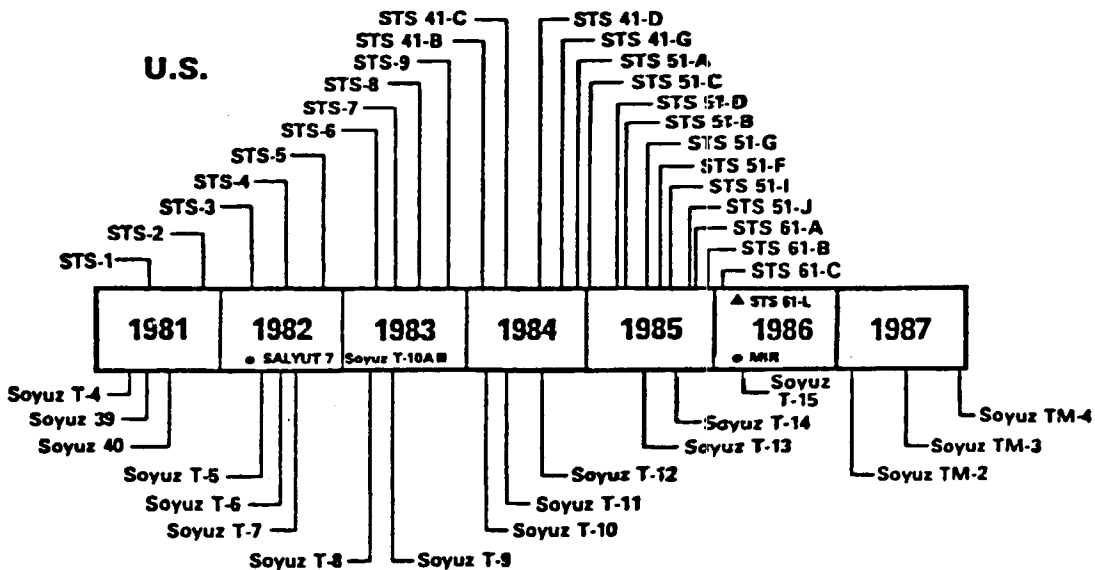


### U.S.S.R.

CONTINUED

## TIME LINE OF U.S. AND U.S.S.R. PILOTED SPACE MISSIONS (continued)

### U.S.



### U.S.S.R.

# ESTIMATED MEAN DAILY INFLIGHT NUTRIENT CONSUMPTION PER PERSON DURING SHUTTLE STS-1 THROUGH STS 61-C

STS Flight	Days #	Crew M F #	RH <sub>2</sub> O <sup>1</sup> gm	NH <sub>2</sub> O <sup>2</sup> gm	KCal	Prot gm	Fat gm	CHO gm	Ca mg	Phos mg	Na mg	K mg	Iron mg	Mg mg	Mn mg	Cu mg	Zn mg
1	2	2			2656	106.8	83.1	358.6	1210	1706	4506	3238	27	387			8
2	2	2	1134	88	1100	58.5	28.0	152.0	687	916	1782	1362	12	154			9
3	8	2	1393	353	1910	66.1	49.6	280.2	885	1210	3010	2244	17	229	1.6	1.9	10
4	7	2	1711	326	2446	85.6	73.5	319.2	954	1474	3506	2558	20	286	2.2	2.2	12
5	5	4	1378	378	2322	73.2	59.7	338.0	931	1464	3333	2415	22	272	2.6	2.6	9
6	5	4	1324	357	1957	75.7	52.1	281.1	787	1227	2829	2226	18	227	1.1	1.5	8
7	6	4 1	1983	281	2535	86.7	76.1	339.4	995	1841	3818	2567	20	336	4.2	3.4	12
8	6	5	1445	378	2517	90.3	67.4	359.3	1026	1729	3697	2822	20	309	2.6	2.4	11
9	9	6	1083	291	1945	68.6	55.7	267.7	833	1382	3138	2393	15	212	1.9	1.9	10
41-B <sup>4</sup>	8	5	1428	411	2684	94.3	81.1	353.4	857	1740	3956	3532	18	355	4.2	3.2	13
41-C	7	5	1687	364	2673	91.5	83.3	344.5	954	1766	3784	3226	16	369	4.0	3.2	13
41-D <sup>4</sup>	6	5 1	1666	382	2143	85.5	63.0	275.8	886	1658	3565	2549	15	303	2.9	2.8	11
41-G	8	5 2	2069	367	2994	103.9	92.1	393.1	1003	1956	4379	3512	19	382	4.7	3.4	13
51-A	8	4 1	1502	283	2383	83.5	73.4	319.8	974	1681	3367	2548	13	273	2.9	2.7	11
51-C <sup>5</sup>	3	5	1937	348	3838	119.7	170.9	395.1	1201	2503	5123	4127	23	543	6.1	3.9	17
51-D <sup>4</sup>	6 7	6 1	1414	282	2219	72.6	74.3	288.6	891	1500	3433	2653	14	315	2.8	1.7	11
51-B <sup>4</sup>	7 7	7	1551	438	2862	90.0	87.6	396.2	968	1761	3698	3684	20	356	3.1	2.4	14
51-G <sup>4</sup>	7 7	6 1	1785	487	3423	107.7	126.1	397.6	1025	2010	4728	4099	19	454	4.1	2.7	19
51-F <sup>4</sup>	7 7	7	1840	363	2783	91.4	91.7	361.2	905	1692	3975	3387	16	362	3.9	2.1	13
51-I <sup>4</sup>	7 7	5	1841	431	2958	105.5	105.2	356.9	980	1779	4547	3630	17	404	4.1	2.3	15
51-J <sup>4</sup>	8 4	5	1240	549	2839	95.4	92.3	378.4	942	1671	3860	3336	16	363	3.8	2.1	13
61-A <sup>4</sup>	7	7 1	1751	418	3287	114.5	116.8	401.5	1224	2008	5654	3517	20	422	4.2	2.3	15
61-B <sup>4</sup>	7 7	6 1	1137	301	2659	93.4	99.2	312.9	806	1484	4244	2785	15	327	3.6	1.9	15
61-C <sup>4</sup>	7 5	7	1772	482	3367	116.1	118.8	411.1	1029	1944	5589	3803	20	431	5.0	2.6	17
Mean (All Flights) <sup>9</sup>			1589	374	2692	92.8	87.8	346.2	957	1715	4037	3132	18	346	3.5	2.5	13
Mean (588 Man Days) <sup>10</sup>			1491	355	2476	85.7	78.0	327.1	911	1606	3673	2913	17	316	3.1	2.4	12
Recommended Minimum Levels					56				800	800	3450	2737	18	350			15
Percent of Calories:					31.0% Prot	54.4% Fat	14.6% CHO										

<sup>1</sup>RH<sub>2</sub>O = Rehydration Water    <sup>2</sup>NH<sub>2</sub>O = Moisture in Food    <sup>3</sup>M = Males    <sup>4</sup>F = Females

<sup>4</sup>Estimated from returned food only. Trash given to Ames and not inventoried.

<sup>5</sup>DOD flight, food was packed for 7 days but they returned after only 3 days. Trash given to Air Force and not inventoried. Fresh food locker not included. <sup>6</sup>Mission extended 2 days. <sup>7</sup>Launch sandwiches not included.

<sup>8</sup>No launch sandwiches sent with STS 51-J. <sup>9</sup>Mean for all flights included 799 man days. <sup>10</sup>Mean omitting flights with unrealistically high caloric intakes i.e., 3000 Kilocalories or more (STS 41-G, 51-C, 51-G, 61-A, 61-C).

**PRODUCTION SYSTEM CHUNKING IN SOAR:  
CASE STUDIES IN AUTOMATED LEARNING**

**Final Report**

**NASA/ASEE Summer Faculty Fellowship Program -- 1988**

**Johnson Space Center**

<b>Prepared By:</b>	<b>Robert Allen, Ph.D.</b>
<b>Academic Rank:</b>	<b>Assistant Professor</b>
<b>University &amp; Department:</b>	<b>Univeristy of Houston Dept. of Mechanical Engineering Houston, TX 77204-4792</b>

**NASA/JSC**

<b>Directorate:</b>	<b>Mission Support</b>
<b>Division:</b>	<b>Mission Planning and Analysis</b>
<b>Branch:</b>	<b>Technology Development and Applications</b>
<b>JSC Colleague:</b>	<b>Robert T. Savely</b>
<b>Date Submitted:</b>	<b>27 July 1988</b>
<b>Contract Number:</b>	<b>NGT 44-005-803</b>

## ABSTRACT

A preliminary study of SOAR, a general intelligent architecture for automated problem solving and learning, is presented. The underlying principles of universal subgoalting and chunking were applied to a simple, yet representative, problem in artificial intelligence. A number of problem space representations were examined and compared. It is concluded that learning is an inherent and beneficial aspect of problem solving. Additional studies are suggested in domains relevant to mission planning, as well as, in aspects related to SOAR itself.

available, the subgoal terminates and pops from the goal stack. The goal stack also serves as the anchor to information in working memory (WM). Each working memory element (WME) is connected to some goal in the stack and can be accessed only by specifying the connection from the goal to the WME via **augmentations**. An example of one such connection is:

```
(goal <g> ^state <s>)  
(state <s> ^binding <b>)  
(binding <b> ^cell c11 ^tile <t1>)
```

where **state** and **binding** are the goal augmentations that provide the connection between the goal <g> and the value of the action attribute. A **chunking mechanism** is provided to summarize the system behavior in terms of subgoals, and also enables the system to learn aspects of problem solving related to subgoals. The overall architecture is presented in Figure 1.

Some of the important characteristics of SOAR are: (1) separation between architecture level and the knowledge level; (2) problem-spaces for representing knowledge; (3) universal subgoalting to resolve impasses ; and (4) production system representation that serves as access paths to information in long term memory. In addition, there is no conflict resolution mechanism in SOAR. All matched productions are fired "in parallel" and add one type of WMEs to the working memory. The process of collecting available information is called the **elaboration phase**. The second type of WME is the **preference element** that is used by the architecture in the **decision cycle** to determine the next goal context. The decision procedure, sketched in Figure 2, controls the elaboration and decision cycles. More detailed descriptions can be found in elsewhere (1-3).

## PROBLEM DOMAIN

SOAR's capabilities are illustrated below in an output trace of the problem-solving process for the "eight-puzzle" problem. The board in the eight-puzzle is represented by nine cells occupied by tiles numbered one through nine with one blank cell. The numbers on the beginning of each line are the decision cycle numbers; elaboration cycles are not shown. The **Build:P** notation signifies a

## INTRODUCTION

SOAR is a production system architecture for a system capable of exhibiting general intelligence. SOAR has three principal characteristics separating it from other architectures. These are: (1) SOAR can be used to solve a range of problems, from routine problems to open-ended problems; (2) SOAR applies a wide range of problem-solving methods required for these tasks; and (3) SOAR learns about aspects of the problem-solving process and is capable of reporting about its performance.

This document summarizes the work performed in implementing simple, yet representative, problems in SOAR. One purpose of this exercise was to be acquainted with the SOAR architecture and implementation. In the course of this work, some general issues were raised and found to coincide with current research topics in SOAR.

This paper is organized as follows. First, a brief description of SOAR is presented. Next, some "toy" AI problems are briefly described and their implementations in SOAR are presented. Penultimately, a comparison of some problems decompositions is examined and the affect of problem presentation on learning and performance is discussed. Finally, future studies that can be performed are recommended.

## SOAR: AN OVERVIEW

SOAR is an architecture for exhibiting general intelligent behavior. SOAR has evolved from a series of production system architectures (1,2). SOAR is embedded in about 255 kilobytes of LISP code, and extends 100 kilobytes of modified OPS5 code.

In SOAR, each task is represented in **problem-spaces**. The problem solving process begins from an initial state, working through the state to subsequent states by applying operators. Stages in the problem solving process are characterized by a goal context, which consists of the current **goal, problem-space, state and operator**. When one stage in the problem-solving process does not have enough information, it creates a subgoal (hence the name "universal subgoaling") to collect the needed information. The new subgoal is then added to the goal stack that keeps track of the goals that were created for this problem-space. When the needed information is

# Soar Architecture

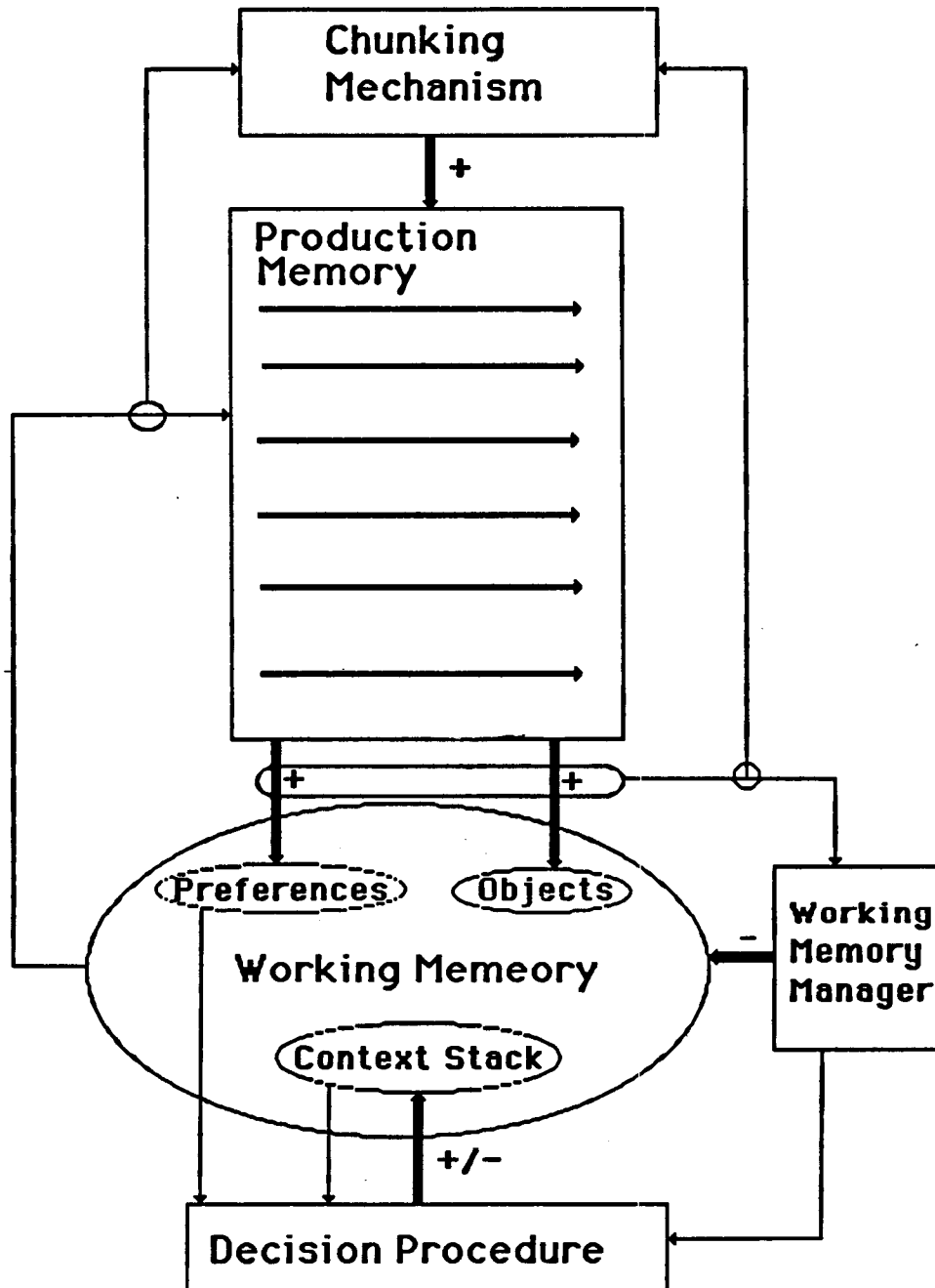


Figure 1

# Decision Procedure

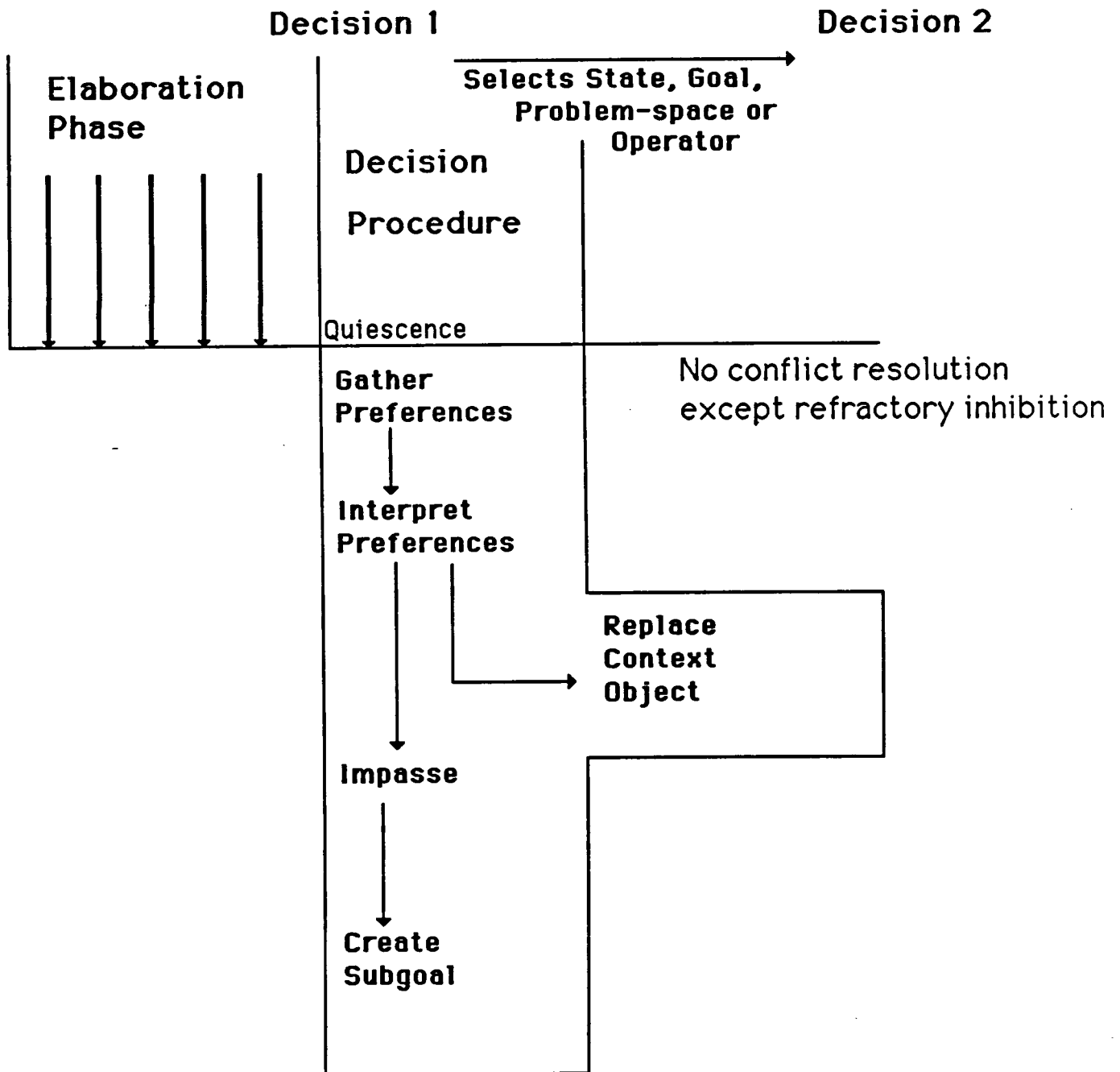


Figure 2

2-4b

new chunk (rule) being created during the trace. The letters G, P, O and S represent the current goal, problem-space, operator and state, respectively.

Learn status: on all-goals print trace

```
0 G: G1
1 P: P3 EIGHT-PUZZLE
2 S: S4
3 O: O34 MOVE-TILE(C22)
```

C22(S4) --> S35

```
-----
| 2 | 8 | 3 |
|---|---|---|
| 1 | 0 | 4 |
|---|---|---|
| 7 | 6 | 5 |
-----
```

```
4 S: S35
5 ==>G: G185 (UNDECIDED OPERATOR TIE)
6 P: P186 SELECTION
7 S: S42
8 O: (O47 O46 O45)
9 ==>G: G190 (EVALUATE-OBJECT (MOVE-TILE(C21)) OPERATOR
NO-CHANGE)
10 P: P3 EIGHT-PUZZLE
11 S: S35
12 O: O40 MOVE-TILE(C12)
```

C21(S35) --> S48

```
-----
| 2 | 0 | 3 |
|---|---|---|
| 1 | 8 | 4 |
|---|---|---|
| 7 | 6 | 5 |
-----
```

C12(S35) --> S53

```
-----  
| 2 | 8 | 3 |  
|---|---|---|  
| 0 | 1 | 4 |  
|---|---|---|  
| 7 | 6 | 5 |  
-----
```

13 S: S53

Duplicate chunk

Build:P200

Duplicate chunk

Build:P202

Build:P204

Build:P205

14 O: O41 MOVE-TILE(C21)

C21(S35) --> S59

```
-----  
| 2 | 0 | 3 |  
|---|---|---|  
| 1 | 8 | 4 |  
|---|---|---|  
| 7 | 6 | 5 |  
-----
```

15 S: S59

16 O: O74 MOVE-TILE(C11)

C11(S59) --> S64

```
-----  
| 0 | 2 | 3 |  
|---|---|---|  
| 1 | 8 | 4 |  
|---|---|---|  
| 7 | 6 | 5 |  
-----
```

17 S: S64  
18 O: O76 MOVE-TILE(C12)

C12(S64) --> S70

```
-----  
| 1 | 2 | 3 |  
|---|---|---|  
| 0 | 8 | 4 |  
|---|---|---|  
| 7 | 6 | 5 |  
-----
```

19 S: S70  
20 O: O80 MOVE-TILE(C22)

C22(S70) --> S77

```
-----  
| 1 | 2 | 3 |  
|---|---|---|  
| 8 | 0 | 4 |  
|---|---|---|  
| 7 | 6 | 5 |  
-----
```

21 S: S77  
goal SOLVE-EIGHT-PUZZLE achieved

>(print-stats)

Soar 4.4 (external release: created April 19, 1987)  
Run statistics on July 26, 1988  
81 productions (1345 / 5703 nodes)  
229.41667 seconds elapsed (43.133335 seconds chunking overhead)  
21 decision cycles (10924.604 ms per cycle)  
46 elaboration cycles (4987.319 ms per cycle)  
(2.1904762 e cycles/d cycle)  
211 production firings (1087.2828 ms per firing)  
(4.5869565 productions in parallel)  
717 RHS actions after initialization (319.96747 ms per action)  
226 mean working memory size (339 maximum, 296 current)  
2890 mean token memory size (10696 maximum, 3893 current)  
22363 total number of tokens added  
18470 total number of tokens removed  
40833 token changes (5.6184134 ms per change)  
(56.7125 changes/action)

The particular implementation of this problem in SOAR requires twelve productions, which can be divided as follows: (1) four rules for setting-up the problem space; (2) one rule to create an operator instantiation; (3) two rules for applying an instantiated operator; (4) one rule for search control; (5) one rule for monitoring states; and (6) three rules for evaluating operators. While the rules described are specific to this problem, encoding other tasks typically involve a similar division of a problem and hence a similar division of rules.

## LEARNING

The chunk below is an example of a rule that becomes part of production memory during the session outlined above.

(SP P218 ELABORATE

(GOAL <G1> ^OPERATOR { <> UNDECIDED <O2> })

(OPERATOR <O2>

^NAME EVALUATE-OBJECT ^ROLE OPERATOR

^SUPERPROBLEM-SPACE <P1> ^OBJECT <O1> ^SUPERSTATE <S1>

^DESIRED <D1> ^EVALUATION <E1>)

(PROBLEM-SPACE <P1> ^NAME EIGHT-PUZZLE)

(OPERATOR <O1>

^NAME MOVE-TILE ^BLANK-CELL <C2> ^TILE-CELL <C1>)

(STATE <S1> ^BLANK-CELL <C2> ^BINDING <B1> { <> <B1> <B2> })

(BINDING <B1> ^CELL <C2>)

(BINDING <B2> ^CELL <C1> ^TILE <T2>)

(DESIRED <D1> ^BINDING <D2>)

(BINDING { <> <B1> <D2> } ^CELL <C2>)

(BINDING <D2> ^TILE <T2>)

-->

(DISPLACED (MAKE EVALUATION <E1> NUMERIC-VALUE 1)

^(NLAM-MAKE (QUOTE (EVALUATION <E1> NUMERIC-VALUE 1))))))

In more transparent form the rule states:

IF any tile is being moved into its final position

THEN the next state will yield an evaluation of 1,

i.e., the next state will be evaluated favorably

This particular rule is one that is created because of the current implementation; the search control rule creates a **worst** preference for those operators not placing tiles in their desired locations. Other search control productions are responsible for the creation of chunks more in line with their method of search control (4).

The search control rules are of importance in the outcome of the problem-solving process. To explore the effect of these rules on the parameters to measure performance, sample runs were performed using three different search control strategies. The results are presented in Table 1. The third column refers to a search control procedure that rejects an operator that will move back a tile into its previous location. The fourth column refers to a search control mechanism that creates a worst preference to a tile that is to be moved out of its desired location. The fifth column corresponds to no

Type of run	Data	Reject * undo	Worst * Preference	No Search
No Learning	number of productions	73	74	82
	elapsed time	177	53	268
	decision cycles	48	21	39
	elaboration cycles	84	43	73
	production firings	442	181	383
Learning	working memory elements	276	301	339
	number of productions	81	77	72
	elapsed time	112 (24)	54 (9)	281 (114)
	decision cycles	30	21	62
	elaboration cycles	55	43	94
Learned	production firings	303	182	572
	working memory elements	323	301	453
	number of productions	81	77	82
	elapsed time	51 (0)	28 (0)	48 (0)
	decision cycles	12	12	12
	elaboration cycles	29	25	29
	production firings	120	104	116
	working memory elements	295	279	291
	chunks used	6	3	7

Table 1: Eight-Puzzle Results

prespecified search control strategy. Each row contains data for each type of run. The data contain the number of productions in the system, the elapsed time to complete the task, the number of decision cycles, the number of elaboration cycles, the number of production firings and the maximum number of working memory elements. The last row, corresponding to runs that have already "learned," indicates the number of chunks that were fired during a problem solving session. The numbers in parentheses indicate the amount of time required for chunking.

As expected, chunking added productions to each system. However, a learned system did not chunk again. Also as expected, a "stronger" search control mechanism (column Four) corresponded to a better overall performance: fewer cycles, fewer rule firings, more efficient memory and more efficient chunking. The reason for this is a *priori* control of the preferences in the "worst\*preference" strategy. Applying specific preference orientations in the search control strategy appears to be an effective way to structure a problem implementation (1,4). Finally, it is noted that even without a specified search control strategy (column Five), SOAR's default strategy is sufficient to solve this problem prior to chunking.

## CONCLUSIONS

A preliminary examination of SOAR has been performed and, from the above-mentioned results, the following conclusions are drawn:

- Learning is a beneficial aspect of automated problem solving in that code can be made more efficient and previously unrecognized knowledge, in the form of chunks, can be created.
- Problem decomposition is the key to an efficient (or even successful implementation). The search control mechanism(s) used in a specific implementation strongly influences the problem-solving process and the learning process.

- Measuring the difficulty of a problem is a nontrivial task. The performance criteria measured by SOAR need to be scrutinized before a definitive measure can be ascertained.

It is clear that additional studies are needed, with more practical problems (such as those presented in (4)), to see if the SOAR architecture can be useful in NASA-related applications.

#### Acknowledgment

The author thanks Lui Wang, Aerospace Engineer for the Artificial Intelligence Section, for his assistance with using the microEXPLORER™ and the Lisp environment.

## References

1. Laird, J.E., "SOAR User's Manual," Version 4, Xerox Palo Alto Research Center, Palo Alto, CA 1986.
2. Laird, J..E., A. Newell and P.S. Rosenbloom, "SOAR: An Architecture for General Intelligence," Artificial Intelligence, 33 (1987)1-64.
3. Laird, J., P. Rosenbloom and A.Newell, "Universal Subgoaling and Chunking, Kluwer Academic Press, 1986.
4. Reich, Y. and S.J. Fenves, "Floor-system Design in SOAR: A Case Study in Learning to Learn," Technical Report EDRC-12-26-88, Carnegie-Mellon University, Pittsburgh, PA 1988.

**N89 - 20061**

**DEVELOPMENT OF A SOFTWARE INTERFACE FOR  
OPTICAL DISK ARCHIVAL STORAGE FOR A NEW  
LIFE SCIENCES FLIGHT EXPERIMENTS COMPUTER**

**FINAL REPORT**

**NASA/ASEE Summer Faculty Fellowship Program -- 1988**

**Johnson Space Center**

Prepared by:	Peter N. Bartram, Ph.D., P.E.
Academic Rank:	Associate Professor
University and Department:	Norwich University Computer Science Engineering Division of Engineering Northfield, Vermont 05663
NASA/JSC	
Directorate:	Space and Life Sciences
Division:	Life Sciences Project Division
Branch:	Project Engineering Branch
JSC Colleague:	Donald Stilwell
Date Submitted:	3 August 1988
Contract Number:	NGT 44-005-803

## ABSTRACT

The current Life Sciences Laboratory Equipment (LSLE) microcomputer for life sciences experiment data acquisition is now obsolete. Among the weaknesses of the current microcomputer are small memory size, relatively slow analog data sampling rates, and the lack of a bulk data storage device. While life science investigators normally prefer data to be transmitted to Earth as it is taken, this is not always possible. No down-link exists for experiments performed in the Shuttle middeck region. One important aspect of a replacement microcomputer is provision for in-flight storage of experimental data.

The Write Once, Read Many (WORM) optical disk was studied because of its high storage density, data integrity, and the availability of a space-qualified unit. In keeping with the goals for a replacement microcomputer based upon commercially available components and standard interfaces, the system studied includes a Small Computer System Interface (SCSI) for interfacing the WORM drive. The system itself is designed around the STD bus, using readily available boards. Configurations examined were (1) master processor board and slave processor board with the SCSI interface, (2) master processor with SCSI interface, (3) master processor with SCSI and Direct Memory Access (DMA), (4) master processor controlling a separate STD bus SCSI board, and (5) master processor controlling a separate STD bus SCSI board with DMA.

Storage times for 512-byte disk sectors ranged from 56 to 22.5 milliseconds without DMA and 53 to 8.9 milliseconds with DMA. The faster times resulted from using large blocks of data per disk write request. With the faster times, the storage rates significantly exceed anticipated data acquisition rates. While DMA is especially attractive, configuring workable systems from available boards without hardware modification is difficult.

## INTRODUCTION

The success of manned space exploration depends in part on an understanding of the physiological effects of microgravity. A major concern of life sciences research in space is to guarantee the health and safety of astronauts for missions of both long and short duration. Additionally, astronauts have experienced many symptoms during exposure to microgravity. These latter symptoms have been subsumed under the rubric of "Space Adaptation Syndrome." It is part of the charter of the Life Sciences Flight Experiments Program (LSFEP) to perform experiments in space that will contribute to the understanding of the physiological effects of space travel on humans and, where possible, to develop strategies for reducing the debilitating effects.

Most flight investigations require the use of a microcomputer for high speed collection, storage, and downlinking of in-flight physiological data. The Life Sciences Laboratory Equipment (LSLE) Microcomputer has served these needs for life sciences investigations in the shuttle Spacelab module. This computer has provided the link between Spacelab experiments and the Science Monitoring Area (SMA) in the Life Sciences Project Division building at the Johnson Space Center (JSC). To perform this function the LSLE microcomputer sends formatted real-time data, in a digital serial form, to a High Rate Multiplexer (HRM) in the Spacelab module, which in turn transmits the data to the ground via a Tracking and Data Relay Satellite System (TDRSS) satellite.

While the old LSLE microcomputer has served its purposes well, it has become obsolescent and must be replaced by a newer design. Shortcomings of the old LSLE microcomputer include a lack of replacement units and parts, an absence of on-board mass storage or real-time data display, an inability to be programmed in a high level language, and poor general specifications when compared to today's standards. Because of these deficiencies, the development of a new LSLE microcomputer is now a high priority item. The requirements for a replacement LSLE microcomputer include all of the data acquisition, HRM downlink, and experiment control functions of the current machine. Additionally, the new microcomputer must have better and more extensive displays, faster data sampling rates, extensive archival storage capabilities for use in middeck experiments where no HRM downlink is available, greater processing speed, more memory, compatibility with popular commercially available systems such as the IBM PC, and the ability to be programmed by the principal investigator using common high-level languages. The design should be based to the maximum extent possible on commercially available boards and should avoid custom hardware as much as

possible. It is hoped that the HRM interface is the only custom design required.

A multiple processor system has several advantages over single processor systems. In the multiple processor system, one processor (called the "master") is reserved to the maximum extent possible for the principle investigator's use, while additional processors (called "slaves") are dedicated to each input/output operation. For example, separate slave processors are used for analog to digital conversion (A/D), digital to analog conversion (D/A), Parallel I/O, Serial I/O, the HRM interface, optical disk control, etc. In this way, work is shared among the many processors. Once this partitioning is defined, separate engineering efforts can attack each task and development work can proceed in parallel. Expansion would have minimal impact on the original parts of the system, as additional processors would handle the additional data acquisition load. This modularization facilitates software and hardware maintenance. It allows on-going replacement of obsolete components at the board level, without significantly disturbing other parts of the system. A chief advantage of this concept is that the modular nature of the system extends into the hardware itself, thereby making development and reconfiguration much simpler.

One of the most important elements of the new LSLE microcomputer is the optical disk. The Write Once Read Many (WORM) optical drive would allow the LSLE microcomputer to be used in the shuttle middeck where no downlink capabilities are available. This will allow the LSLE microcomputer to be used on a much larger number of flights (including those not flying the Spacelab), thereby greatly expanding opportunities for the Life Sciences Flight Experiments Program. In most cases, the investigator will prefer real-time data downlink, but as flight opportunities are limited, he may choose to fly his experiment in the middeck rather than to risk losing the opportunity altogether. In other cases, the investigator may be given a late opportunity to add his experiment to an existing fully developed payload. In this case, time may not permit the development of ground software to handle the HRM data stream. Again, the use of an optical disk will reduce data handling costs and development times. It has been estimated that the cost of ground software to handle the HRM data stream for life sciences experiments is \$250,000 per mission, regardless of the number of experiments served. The use of the optical disk instead of the HRM would remove this cost entirely. In summary, the use of an optical disk will allow single experiments to be inserted into missions on relatively short notice and will permit full use of the middeck by life science investigators. Other bulk storage systems, such as streaming tape and other magnetic devices, do not offer the robustness and high data density of the optical disk.

The objective of my work this summer was to develop prototype software needed to store collected data on an optical disk as well as to recover it. Also, timing studies were to be performed to determine if data storage rates would

be sufficient for the data collection sample frequency required of the new system.

## EQUIPMENT USED

The commonly used STD bus was chosen for the system backplane. This bus is well supported, with several hundred manufacturers offering several thousand boards. Thus, prospects are excellent for future upgrades in keeping with advances in technology.

The STD bus processor boards selected for the system are based on the Intel 8088 and 80188 processors. Advantages of these boards include processor compatibility with the familiar IBM PC family and the availability of a wide range of software and hardware support products. In the system under study, the master processor board is the Ziatech Corporation ZT8815, containing an 8 MHz 80188. This board has computing power similar to the IBM PC/AT. Of particular interest, the 80188 processor chip contains Direct Memory Access (DMA), which provides fast memory -- peripheral data transfer in parallel with other processor functions. Additional processor boards, slave processors, are Ziatech ZT8830 boards. These contain 8 MHz 8088 processors, with computing power approximating that of the IBM PC. Both master and slave boards contain a standard SBX interface and connector for "piggyback" peripheral modules. Hundreds of modules are available from dozens of vendors, providing real-time input and output hardware. Thus self-contained data acquisition subsystems can be formed using, for example, an SBX analog to digital converter attached to a slave processor board.

The optical disk interface is the Small Computer System Interface (SCSI) standard. One configuration uses a slave processor with the Zendex Corporation model ZBX-280 SCSI SBX module. The ZBX-280 is based on the NCR 5380 SCSI controller chip. The same SBX module attached to the master processor is an alternative configuration, which allows the use of DMA, not available with the slave processor boards. Another alternative SCSI interface is the Ziatech STD bus peripheral board, the ZT8850, which provides a chip set for SCSI as well as other disk control devices. The ZT8850 includes its own DMA. Compared to the Zendex ZBX-280, the ZT8850 is simpler to program, with more SCSI signals generated by hardware rather than under program control. The master processor board controls the ZT8850 as a bus peripheral, unlike the ZBX-280 which is directly interfaced and does not reside on the STD bus. With the ZT8815 processor and the ZT8850 SCSI boards, DMA may be implemented using either the 80188 DMA of the ZT8815, or the ZT8850 on-board DMA circuits.

The optical disk drive is the Optotech model 5984. This drive was selected because Mountain Optech (Boulder, CO) makes a space qualified version of the drive (model 200SES) under contract to Goddard Space Flight Center. The model 5984 is identical in electrical and software characteristics to the model 200SES, but is less expensive. Both are 200 megabyte (per side) capacity optical disk drives, using 5 1/4 inch removable media cartridges.

A summary of the equipment used and location of principal vendors is given in Table 1.

### SOFTWARE DEVELOPED

Test programs were written to assure that the optical disk could be both written and read in the various configurations considered, and to determine read and write times. Best case conditions were used in that no other competing activity was required of the processor chip or, where used, the STD bus. In all cases, a main program in C calls a collection of assembly language functions. For support of the Zendex SCSI piggyback (NCR 5380 chip) interface with either master or slave processor board, the assembly language routines are revisions of routines provided by Mountain Optech. Each principal SCSI phase is handled by a separate call: reset, selection (including arbitration), status, command, message in, data output, and data input. (The message out phase has not been implemented.) The data transfer is byte by byte, with polling and handshake signal generation by software. The calling program can thus check for proper progression of phases and handle error conditions. Additional functions were written for DMA use for the data input and data output phases with the Zendex SCSI interface used with the ZT8815 master processor board.

A separate set of assembly language functions, matching the names and organization of the above routines, has been prepared for use with the master ZT8815 board and ZT8850 peripheral SCSI board. These require no significant change in the C calling programs written for other configurations. While the example code available from Ziotech for the ZT8850 support was influential, these routines followed the organization of the functions for the Zendex interface, rather than the task oriented Ziotech routines. The routines for the ZT8850 are generally simpler than those for the NCR 5380 based Zendex interface. The ZT8850 does not support arbitration. The handshake signals required for data transfer are hardware

**TABLE 1. -- EQUIPMENT USED.**

**HARDWARE**

Ziatech 8862 Card Cage and Power Supply (STD bus)  
Ziatech ZT8830 Intelligent I/O Control Processor (slave)  
Ziatech ZT8815 80188 based CPU card (master)  
Zendex ZBX-H280 SCSI controller multimodule (iSBX)  
IBM PC compatible personal computer for downloading programs to STD bus boards  
Digital storage oscilloscope for obtaining timing data  
Mountain Optech Model 5984 Optical disk drive

**SOFTWARE**

Ziatech 8830 Debug software (and ROMs) to load and debug programs  
Ziatech 8815 Debug software  
Microsoft C Language, version 4.00  
Microsoft Macro Assembler, version 4.00  
Microsoft Linker (loader), version 3.51  
Microsoft MS-DOS operating system, version 3.10 (for program development)

**VENDORS**

Ziatech Corporation  
3433 Roberto Court  
San Luis Obispo, CA  
93401

(805) 541-0488

Zendex Corporation  
6700 Sierra Lane  
Dublin, CA  
94568

(415) 828-3000

Mountain Optech  
2830 Wilderness Place  
Suite F  
Boulder, CO

80301

(303) 444-2851

Microsoft Corporation  
10700 Northrup Way  
Bellevue, WA  
98004

(206) 882-8089

generated. The data transfer is still byte by byte, with handshake signal polling to avoid data loss. Some DMA support routines for the ZT8850 have been developed.

Copies of software developed for this project may be obtained from Peter N. Bartram, Division of Engineering, Norwich University, Northfield, Vermont 05663 (telephone 802/485-2263), or Donald Stilwell, NASA/Johnson Space Center, Mail Code: SE3, Houston, Texas 77058 (telephone 713/483-7308).

## RESULTS

Polled data transfer can be performed in all configurations studied (ZT8830 + ZBX280, ZT8815 + ZBX280, and ZT8815 + ZT8850). DMA writing to the optical disk may be performed using the ZT8815 with ZBX280 configuration, provided the SBX signal TDMA, not implemented by the ZT8815, is grounded (requiring the addition of a wire on either the ZT8815 or the ZBX280). The ZBX280 uses this signal for one alternative for ending DMA transfer. With it left floating, as on the ZT8815, the ZBX280 attempts to halt DMA prematurely. Even with this change, with which DMA writing to disk works well, DMA reading of the disk returns incorrect values. The source of difficulty has not been determined with confidence. The ZT8850 DMA controller functions correctly, provided the memory used is not on the ZT8815 processor card controlling it. (A separate memory board was used.) Also at the time of this writing, code for using the ZT8815 processor DMA with the ZT8850 is under preparation.

For polled data transfers, timing results were similar for both reading and writing. For the LSLE replacement microcomputer, disk writing in real-time is critical, whereas reading will be performed at a later time, when speed is of lesser importance. With the ZBX280 SCSI interface, writing a single 512 byte sector each call for data output, the average write time was less than 56 milliseconds per block, regardless of processor. With each write request specifying a ten-sector block (5120 bytes), the average write time for ten sectors was 250.4 milliseconds (25.04 milliseconds per sector). Some ten sector blocks required 225 milliseconds, others 276 milliseconds. These times were identical for both the ZT8830 and ZT8815 processor boards. For write request block sizes of 25 sectors (12800 bytes), the average time required was 562.5 milliseconds per block (22.5 milliseconds per 512 byte sector). This is over 22 kilobytes per second. Since each real-time measurement results in a two-byte value, data storage in excess of 11000 samples per second is possible.

For the ZT8850 SCSI with no DMA, write times for single sector requests averaged 52.7 milliseconds per sector. For ten sector blocks, the average time per block was 197 milliseconds (19.7 per 512 byte sector). With 25 sector blocks, the time averaged 369 milliseconds (14.8 milliseconds per 512 byte sector). With 25 sector (12800 byte) blocks, over 33.8 kilobytes per second can be stored, or over 16,900 measurement samples per second.

With DMA using the ZT8815 with ZBX280, single sector write times averaged 52.7 milliseconds per sector. For DMA using blocks of 10 sectors per write request, the write time per block fluctuated between 124 and 176 milliseconds, averaging less than 150 milliseconds per block (15 milliseconds per 512 byte sector). The time required to write blocks of 25 sectors fluctuated between 210 and 260 milliseconds, averaging 221 milliseconds per 12800 byte block, or 8.84 milliseconds per 512 byte sector. Thus with DMA, storage rates in excess of 56 kilobytes per second (28000 measurement samples per second) are possible.

Using the ZT8850 STD bus SCSI board with DMA (as provided on the ZT8850 board), single sector write times also averaged 52.7 milliseconds per sector. Average write times using blocks of ten sectors per write request also was less than 150 milliseconds per block, 15 milliseconds per sector, and the same fluctuations in time were observed. However, the times for writing blocks of 25 sectors were longer: fluctuating between 260 and 312 milliseconds, averaging about 269 milliseconds per block. With this, about 46.4 kilobytes per second average storage rate results.

## CONCLUSIONS AND RECOMMENDATIONS

Without DMA, data storage rates are marginally adequate for anticipated data collection rates. With DMA, storage rates exceed any envisioned requirements. DMA is particularly attractive because during data transfer to the disk, the processor is freed for other tasks.

At this point, it appears better to use the master processor for control of the SCSI interface to the optical disk. At the time of this writing, no slave processor supporting DMA is available with an SBX piggyback connector (for an SCSI interface), though these are under development. In order to gain the benefits of DMA, the master must be used to control the SCSI interface. Even if the slower transfer rates of currently available slave processor boards supporting the SBX piggyback modules (such as the ZBX-280 SCSI) are accepted, significant master processor usage would still be involved. This is because one slave cannot access memory of others. Thus

the master would have to transfer data from data acquisition slaves to the SCSI slave memory. With the master controlling the SCSI interface, no memory to memory copy is required. The master can send data to the SCSI directly from data acquisition slave board memory.

Block size is the most important factor for write speed. This is probably because with large numbers of contiguous sectors written with one write command, fewer disk seeks are required to properly position the write heads. DMA only marginally improved single sector block write times, where seek time seems to be limiting. With larger block sizes, not only is speed improved with both methods, but DMA gives markedly improved performance. With large amounts of data being written with each output command, data transfer time becomes limiting, as the number of seeks is reduced. The uneven write times can probably be attributed to the occasional write head repositioning required for two contiguously addressed sectors located on adjacent disk tracks.

However, DMA is not without difficulty, especially in mixed-vendor configurations. The TDMA signal incompatibility has been mentioned earlier. It is suspected that other problems of non-standard standards may be involved in using DMA in other configurations. The system designer must have a good understanding of both hardware and systems programming.

Several recommendations now can be made. (1) Further work is needed to test alternative DMA configurations. (2) As the block sizes were here chosen arbitrarily, optimization of the number of sectors written per write request in light of memory requirements is required. Along with this, the need to interleave data from several data acquisition slaves needs to be considered in determining block size. (3) It is important that the timing studies be repeated in a prototype system with all parts functioning. With other activity on the STD bus and more demands on the master processor, it should be verified that data storage rates will be maintained at a level exceeding the requirements of the data acquisition slave processors. (4) When slave processor boards supporting DMA and the SBX piggyback modules become available, the wisdom of placing the SCSI under control of the master processor should be reconsidered.

COMPUTER TECHNOLOGIES AND INSTITUTIONAL MEMORY

Final Report

NASA/ASEE Summer Faculty Fellowship Program-1988

Johnson Space Center

Prepared By:	Christopher Bell
Academic Rank:	Assistant Professor
University & Department:	Illinois State University Department: Applied Computer Science Normal, Illinois 61761-6901
Prepared By:	Roy Lachman, Ph.D.
Academic Rank:	Professor of Psychology
University & Department:	University of Houston Department: Psychology Houston, Texas 77004
NASA/JSC	
Directorate:	Space and Life Sciences
Division:	Man-Systems Division
Branch:	Crew Station Branch
JSC Colleague:	Barbara Woolford, MA Frances E. Mount, MS
Date Submitted:	August 12, 1988
Contract Number:	NGT 44-005-803

## ABSTRACT

NASA programs for manned space flight are in their 27th year. Scientists and engineers who worked continuously on the development of aerospace technology during that period are approaching retirement. The resulting loss to the organization will be considerable. Although this problem is general to the NASA community, the problem was explored in terms of the institutional memory and technical expertise of a single individual in the Man-Systems Division. The main domain of the expert was spacecraft lighting, which became the subject area for analysis in these studies. The report starts with an analysis of the cumulative expertise and institutional memory of technical employees of organizations such as NASA. A set of solutions to this problem are examined and found inadequate. Two solutions were investigated at length: hypertext and expert systems. Illustrative examples were provided of hypertext and expert system representation of spacecraft lighting. These computer technologies can be used to ameliorate the problem of the loss of invaluable personnel.

## DESCRIPTION OF THE PROBLEM\*

Personnel represent a major share of the resources of any organization. As these individuals interact with others in and outside of their organization, as problems are confronted and solutions identified, the organization's personnel increase in value to the organization. The solutions, the interactions, the actions of the personnel of the organization during their association with it contribute to the institutional memory of the organization. Each employee is privy to part of this memory to the extent the employee has participated in the organization's activities. Much of this institutional memory resides only within the employees. As such it is at risk with the employee. When any employee resigns, retires or dies, their unique part of the institutional memory is lost.

This paper examines this problem as it confronts NASA. It describes the early stages of an effort - using the tools of expert system development and hypertext data management schemes - at NASA-Johnson Space Center/Life Sciences Directorate to capture some of this institutional resource before it is lost. While NASA is a unique organization in many respects, its efforts to confront and resolve this problem may provide some guidance for other organizations which are encountering similar problems.

### Loss of Institutional Memory

Institutional memory is the history of the activities and interactions which have occurred within an organization. Some of this history is recorded in the form of memos, papers, research studies and other archival material, but a large share is resident in the ephemeral memories of employees of the organization. Much of the information is common to more than one individual, but each individual has a unique combination of elements of the larger whole. They can also claim some unique elements of the institutional memory. As these individuals are separated from the organization - through transfers, retirements or death - the unique components of the institutional memory which they possess are lost to the organization.

Not all individuals within the organization can be considered to possess equally valuable components of the institutional memory. The corporate technical experts

\*The order of authorship is alphabetic: C. Bell, specialized in the hypertext solution, R. Lachman in expert systems.

will be more highly valued than the corporate chauffeurs. Ideally, a process for the identification of the critically important experts and the capture of their critical experiences can be in place early in the expert's association with the organization. Such a process would greatly ease the problems of retaining within the organization the expert's uniquely valuable knowledge. The failure to identify these individuals prior to their leaving the organization will ensure loss of their unique components of the institutional memory.

NASA has a workforce which is rapidly approaching retirement age. Due to the unfortunate conjunction of similar hiring dates - associated with the start of the space program - and the number of individuals rolling over their military retirement into the federal system, some divisions ( e.g.; Man-Systems Division ) have over 50% of their technical and managerial personnel eligible for retirement. This is an aging population as well, subject to the impact of degenerative disease processes and common mortality risks.

Senior experts serve other roles in organizations beside those related to the application of their expertise and as a repository of institutional memory. They provide mentor services to junior personnel, acquainting new hires with successful and unsuccessful experimental and behavioral procedures. With the loss of the senior individuals the efficiency of the organization declines as non-productive techniques are retried, as mistakes are repeated by the less experienced new personnel.

#### Approaches to Solution of Problem

Several techniques have been tried for managing the problems associated with the loss of senior personnel. These can be grouped into those designed to reduce the rate of personnel loss and those designed to compensate for this loss. The organization can attempt to reduce the number of personnel lost to retirement by providing compensation - increased pay, more recognition, etc. - to these individuals. Appeals to corporate loyalty or patriotism may be tried to retain vital personnel. Legal adjustments - such as removal of mandatory retirement laws - have reduced arbitrary barriers. If retirement is not preventable, efforts can be made to reduce the impact of the expert's leaving. These efforts might take the form of long term post-retirement consulting assignments. A phased retirement package made of gradually extended vacation time allotments during the employee's last few years of service might help to reduce the impact to the organization of

sudden retirement.

Several approaches are available to compensate for the loss of personnel. To try to retain some elements of expertise at risk, an organization may assign apprentices to senior experts. These apprentices can reduce the impact of the loss of the expert by providing a partial replacement during a transitional period. New archival technologies are being introduced which will provide computer assistance to the problem. The organization can manage the problem in the traditional fashion of accepting the loss of the expert and their unique institutional memory components.

None of the retirement-rate reduction programs described compensate for the loss of experts due to age related or accidental processes. Consulting programs or phased retirement programs are weakened by two conflicting issues. To the extent that the expert is out of the normal flow of the organization's information and duties - referred to as the "loop" - the value of the expert's unique contribution is diminished. The more attempts are made to keep the expert in the "loop" the less others in the organization are able to accept and begin to compensate for the loss of the expert. There is a morale problem associated with excessively delayed retirement programs as well. The senior expert should be able to enjoy a healthy retirement period. Junior employees may begin to resent the fact that advancement routes are being blocked by the continued presence of the senior experts.

The use of apprentices, trainees or proteges would seem to give a semblance of continuity to activities requiring high levels of expertise. The success of this approach is dependent upon several factors. It assumes that there are interested trainees available. It assumes that the expert is sufficiently introspective to identify the critical features of themselves that make them valuable to the organization and is able to determine effective techniques to transfer these features to another. The expert needs to be able to verbalize his expertise and identify situations where it may be safely practiced by the trainee. These are not common skills nor are they easily learned. Finally it is unlikely that the expert can pass along the authority and knowledge associated with many years experience in a short time period.

Simple acceptance of continuing losses is the most commonly chosen alternative. Organizations such as NASA will find the cost of these losses very onerous. This leaves NASA with a strong motivation to explore

alternate archive generating and organizing procedures or other technological fixes.

An active archival program with a professional staff to gather, cross-reference and maintain significant documents is an expensive option. A difficulty in using such an approach for the problem of the loss of institutional memory through personnel attrition is that the material which is retained is the material which is retainable - printed text - and not the critical material of attempts successful or not, of judgements validated or disabused. These are retained by the expert and make up a crucial portion of the expert's unique component of the institutional memory. While these experiential items may be vital to the expert and their handling of new problems, they are not nearly as likely to appear in archival storage as memos or minutes of informational meetings. Had the expert maintained an active report or diary of problem solutions and attempted strategies, of backgrounds to decisions made and processes observed, the problems associated with both the archival and the apprenticeship approaches would be reduced. If the retiring expert can be induced to create a set of documents detailing the significant issues used to identify and solve problems, the thought processes and attack strategies needed, then the problems outlined in the preceding paragraphs will recede. It is unfortunate that most experts find it extremely difficult to accurately verbalize the techniques they use to identify solutions. Most are not even conscious of the procedures they are applying to the solution. Memory that may be critical to the decision making process may not be consciously available until needed for the solution of a particular problem. The reminiscence approach has a further negative associated with it. If a particular domain or pattern of problems has not been addressed recently, the information may not be brought to conscious memory in time to be added to the verbal record.

Even with complete memoirs of retirees, problems of organization, selection, access and structure remain. Access procedures must be simple enough to facilitate their use by remaining employees. They need to be robust enough to survive and be able to produce useful information for users who may not have the time to become facile in using a sophisticated data search package. Cross references between memoirs and other archival material need to be maintained, and a data dictionary of terms and synonyms established. The standard to work toward is an on-line system available to users within their desk environment or through simple telephone connections - equivalent in ease to contacting the former employee. It can be anticipated that a significant

portion of the material provided will be redundant, but a redundancy of information is to be preferred to a shortage of information. A winnowing process can control for information duplicated by other retirees or available from other sources.

Experts serve as more than information repositories. They achieve their status through the appropriate application of their knowledge to the situations they confront - they make decisions. While present software and theoretical activities point toward efforts to duplicate the broad decision making processes of experts, the most successful applications thus far have been those where the domain of interest is very narrow. Expert system tools have been most successfully applied in highly bounded situations, where the range of variables of interest can be anticipated and their interactions predicted and encoded. Unbound situations provide the greatest challenge to the extension of this technology. The problem of the loss of institutional memory and consultative resources is a very unbounded situation.

The replication of the tasks of the expert may be approached as a problem in artificial intelligence (i.e. as an expert system problem), or as a data base/archival problem. Given the specific needs of the person seeking to use the expert as a consultant, a traditional query language/key word - synonym archival approach will require substantial knowledge of the information domain prior to effective use. Many of the individuals contacting an expert express great unwillingness to become more familiar with the knowledge domain of the expert than they already were. Too often keyword searches result in too many "hits" to be useful to a simple informational search or result in too few "hits" because of a lack of overlap between the user's vocabulary and that of the expert. Scrolling is a very inefficient way to search a data base.

Given the difficulties outlined above for using standard archival approaches, the need for alternative approaches seems evident. We will examine two: hypertext and expert systems.

#### HYPERTEXT

If sophisticated data storage systems could be combined with easy to use linking or searching schemas, many of the difficulties of recovering information from a text base could be reduced. Such a combination of linked knowledge stores and an effective front end interface is available in present software shells under the generic

term hypertext or hypermedia.

The concept of hyper-text is quite simple: Windows on the screen are associated with objects in a (knowledge) base, and links are provided between these objects. (Conklin, 1987).

### General Description

The term hypertext is used in this paper to describe a set of techniques for creating and linking text-oriented nodes of information which may be accessed dynamically during any given trial. Each node represents an item of interest in the hypertext knowledge base. Links are established to other nodes containing information associated with the node. Links are conceptual pathways that are used to move from one node to another. A user browses in the hypertext knowledge base seeking information. This description of hypertext is meant to be general. Each experimental or commercial shell or developed package will differ depending upon the creator's personal view of the relative importance of various features. For an excellent review of the specific implementations of hypertext systems, see Conklin (1987).

Jumps between related nodes occur at the user's discretion. The user has the option of pursuing any line of inquiry by following links to nodes that may contain relevant information. While the links need to be pre-established, the exact path chosen through the hypertext knowledge base is at the user's discretion and thus not rigid. Links may be followed in both directions. Any given node may serve as the target for one or more links and may have no, one or several links leading from it. Frequently, a graph or map of the hypertext knowledge base is maintained by the system to serve the user as a graphic memory aid. Whether the underlying linkage structure should be hierarchical or not seems to be at the developers' discretion. Conklin (1987) presents arguments for both approaches. Although most users inherently favor a hierarchical structure some need for the ability to jump between structures seems to be supported. Users wishing to jump from one node to another may do so with minimal keystrokes. The software keeps track of the user's location and efficiently permits the user to quickly leave a node to move to a higher level (assuming a hierarchical structure) or to another node with ease. The simplest analogy to a large non-structured hypertext system is to visualize interlaced cobwebs with nodes of information at each of the cross points. The user is like the spider running along the web, each strand representing a link. This is in

contrast to traditional data base schemas where the pattern of linkage and of searching is rigidly structured (as in hierarchical and network designs) or designed for pattern matching (as in relational data base organizations).

#### Extraction of nodes

The creation of the hypertext oriented system requires the efficient extraction of nodes from the background material. Nodes are edited to represent the information they contain in simple useful text. Text quantity may correspond to a sentence, a paragraph or a screen. Given the desire to produce useful and frequently used reference materials nodes should be edited with a bias favoring the simple and the short over the long and convoluted. As the nodes are extracted, they are linked to related nodes. These links are established to provide possible routes of access for users. While it is impossible to anticipate all possible routes of user interest, it is suggested that the original source of the material might serve very nicely as an initial approximation of the pattern of relationships. Cross links are then made to relevant topics in nodes extracted from other documents. A cognitively manageable structure needs to be retained. Commonly, clues to links to other nodes have been in the form of embedded references within text or menu lists (see Fig. 1 and Fig. 2). Either the first line of the node is given or a node title is provided to aid the user in deciding whether to use a provided link to the node. Both approaches have support in commercial systems.

#### THE HYPERTEXT PROJECT

The role of the expert is both as a consultant and as a decision-maker. As a consultant the expert is able to bring together information from previous work which may be relevant to the project at hand. The decision-making aspect involves using this information to reach a conclusion about the problem. The information management aspect of the expert can be approximated using hyper-text oriented hypertext knowledge base procedures. Information from the expert's descriptions of previous projects can be merged with data from other sources to approximate the knowledge of the expert, and then linked to provide a user with access to the information of the expert. Suggested sources of material to put into the hypertext knowledge base would include retirees' observations, technical reports and other collateral material. This hypertext knowledge base would be tied to a user

ORIGINAL PAGE IS  
OF POOR QUALITY

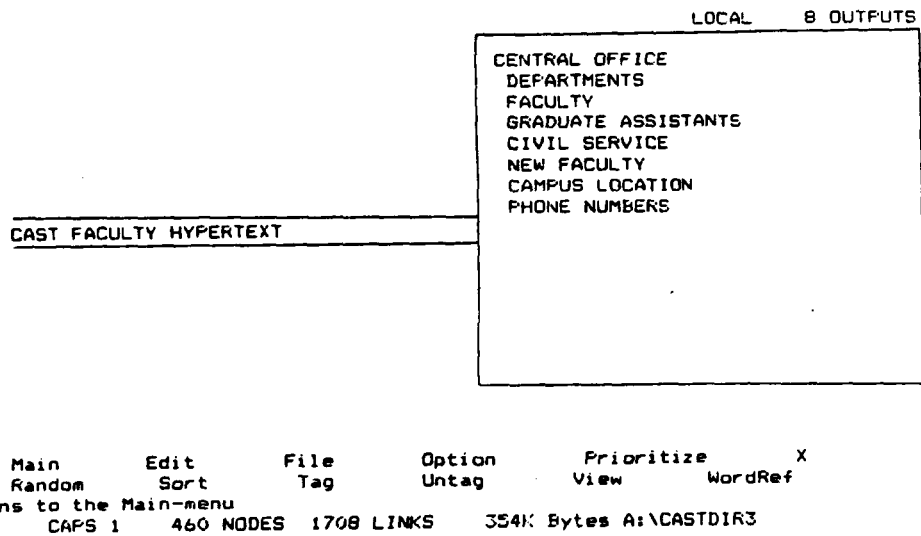


Figure 1. - Menu oriented hypertext screen.

to users the actual methods that will lead to a particular decision. With hypertext, users directly participate in each and every decision that leads to a particular expertise. This process of openly displaying the structure <FILE75 STRUCTURE> and uses of knowledge readily transmits it to users of hypertext systems.

**SENSITIVITY ANALYSIS** Experts machines usually provide a single answer supported perhaps by a confidence factor such as 82 percent certainty (whatever that means). Hypertext systems allow users to rapidly test alternative paths <FILE56 BROWSE> to see how sensitive the advice may be to changes in the initial assumptions.

In my opinion, given the years of AI promises and the relative lack of success in creation and delivery of workable experts systems, the most important pocketbooks are rapidly closing against further use of computer processing of rule-based systems as a method of vending expertise from a disk.

Generally, over the last 30 years, technologists haven't made operations research practical. Consequently, if you can't build machines that are effective in quantitative reasoning, how can you build machines that are effective in subjective reasoning. Making subjective decision machines work is at least several magnitudes more difficult. <FILE27 REASONS>

-- More -- Page 4 Columns 0 - 79 Rmargin: 77 FILENAME: file21

Figure 2. - Embedded link hypertext screen.

interface designed for easy, effective interactions. With the passage of time the organization will progress beyond the areas of the technical competency of the retired expert. The number of archival inquiries can be expected to drop. The nature of those inquiries can be expected to change from specific information seeking to general viewpoint or attack strategy issues. This represents the best presently available solution to the problems associated with the need for a partial replacement for retiring experts.

The expert's interpretation of the information cannot be duplicated, nor can the expert's judgement as to the relative importance to attach to the various nodes. The user still must make a judgement as to the weight to put on the information so presented. The advantage to the user is that the information is available quickly. Associated information is rapidly accessed and relative importance to the project at hand can be judged. The user is integrated with the computer hypertext knowledge base, rapidly pursuing leads and identifying interesting nodes and rejecting those of little interest. While the expert is not available, the user is able to make approximation of the expert by accessing the same information, linked in the same manner as the expert would have linked it.

NASA has a critical need to establish protocols for capturing the expertise of retiring expert personnel. To this end an individual was identified who has served as a consultant in a variety of areas including lighting and window design. A brief sampling of the telephone calls received by this individual supported the claim of being consulted over fifty times per month on issues related to areas of expertise. This particular individual is also very verbal and it was hoped that this would also help, given the prototype nature of the project.

#### Procedure for Gathering Information

Initial data was gathered during interviews with the subject. The interviewers quickly discovered that it was necessary to isolate the subject from his normal work environment as the demands on the subject for consultation were frequent. Interviews were transcribed and the transcription files served as the core for the hypertext knowledge base.

Initial interviews were minimally directed. The subject was asked to discuss his experiences with lighting issues during the development of various manned space efforts. Review of the interviews identified lacunae which were used as prompting points in

subsequent interview sessions. There were two objectives being met during the interview series. One was to provide a core of material for the development of the hypertext knowledge base. The interviews also served as a means to isolate some examples of decision making that could serve as models during the development of the proto-type expert system. Both interviewers were present during the sessions. Some conflicts developed as to directions for follow-up questions or deciding the point at which a narrative needed to be redirected. Sufficient information was gathered for both efforts.

#### Development of Hypertext Knowledge Base

Given the goal of creating a useful and used product, the hypertext knowledge base uses the narrative material as one of several important sources. The second major source is a NASA document titled Man-Systems Integration Standard. This document is a product of the Man-Systems Division and serves as a summary of human factors related information for NASA. Other sources useful to this project are published reference material dealing with vision and light. Material from the latter sources can be updated as new interpretations of data cause changes to be made in the scientific publications relevant to the topic area.

The hypertext document being produced for NASA uses the commercial hypertext software shell HOUDINI. Efforts are being directed toward examining issues related to determining the most effective interface for NASA personnel and other technically oriented users.

Questions to be addressed include:

How much freedom of movement - cross-linkages as opposed to hierarchical links - is needed by the casual user in contrast to the experienced user?

Should the user be able to annotate the hypertext knowledge base?

Should the user be able to add new links or nodes?

How does training effect user format preferences?

These will be discussed in future papers by the authors.

Users of hypertext systems have noted some operational problems. There is a significant cognitive overhead involved in keeping track of the information sought and in identifying nodes of immediately useful information giving a large number of irrelevant or partially relevant nodes. User disorientation - lost in hyperspace -

occurs when the user loses their cognitive map of the relationships between nodes. There is the risk of a combinatorial explosion of links if nodes are thoroughly cross-linked. Each of these areas will be actively addressed. Hypertext is not the final solution to the problems associated with the loss of institutional memory. It does present a methodology for reducing the effect to the organization of the loss of institutional memory through mortality of corporate experts. This is an expensive loss. Whether modern organizations can acknowledge the threat of this loss and make the required investment of time and effort to begin to compensate for it, is another question.

## EXPERT SYSTEM APPROACHES

### Project Scope and the Expert System Life Cycle

In the development of expert systems, the most important judgement is the one that determines the suitability of a task for expert system support. A positive decision in this regard represents the first in a series of sequential steps that are recommended for the development of an expert system prototype prior to the construction of a full-scale knowledge-based system (Waterman, 1986). Prototype development is one phase of the expert system life cycle. The life cycle consists of five major phases: system conceptualization and formulation of requirements, selection of software tools, developing the prototype, constructing the end product, and maintaining the product.

The summer project, according to its initial conception, revolved around the professional activities of the Man-Systems Division internal specialist in spacecraft lighting and windows, CW. Since CW is approaching retirement, it appeared prudent to examine the options available to NASA in any effort designed to preserve the operational experience he has accumulated during a period of about 30 years service to the organization. The expert system approach was among the two options that were selected for detailed examination in consultation with NASA colleagues. The scope of the project was conceived as more in the nature of a problem analysis and overview than the construction of an expert system prototype. Consequently, this part of the project is best characterized as a pre-prototype problem evaluation and the initial phase of prototype construction.

Typical knowledge-engineering methodology was used. The flexibility of the method is an asset but the concomitant absence of standards can be a serious impediment to the development of a knowledge-based system. Logically, the first step in the methodology consists of an evaluation of

the expertise of the domain expert selected for the project. It is essential to initially determine if the candidate consultant is in fact a genuine expert with respect to the subject domain. However, a necessary prerequisite for that judgement, as well as subsequent steps in prototype development, is an adequate level of familiarity with the domain of knowledge. Consequently several weeks were spent examining primary sources in lighting engineering and design (e.g. Kaufman and Christensen, 1984; 1987). The initial acquisition of domain vocabulary and conceptual structure was later evaluated, corrected, and extended by questions posed to CW during formal knowledge engineering sessions. During the period devoted to the study of the literature on illumination engineering, a log was kept of the consultations that the expert provided to other NASA professionals, the meetings that he attended, and the tests that he conducted.

The consultations are best characterized as problem solving sessions and covered a surprisingly broad range of topics. The expert devoted considerable time to analyzing competing designs for the cupola windows in the space station. In addition to forecasting general visual effects of window design modifications, he constructed AITOFF Equal Area Projection of the Sphere showing the kinds of visual function (foveal vision, color vision, binocular, etc) that could be lost as a result of several contemplated changes in window design. He also gave technical advice on attributes of space station windows such as size, number of panes, shape, replaceability, construction materials and selection criteria. The second area consisted of advice on space station vibroacoustics, including issues of subliminal vibration. Third, he developed TV overlays for viewing the grapple on the Canadian arm for handling payload during STS-24. Fourth, he conducted ambient lighting experiments that compared the relative efficacy of LCD and gas plasma displays for the on board laptop computers that display the location of the Orbiter. Fifth, he presented arguments dealing with the design of central lighting for workstation areas on the space station. (Someone wanted to change the overhead lighting specifications from 4800K to 3200K ostensibly to achieve superior resolution for TV monitoring of workstation activity. CW had photos in his files from previous tests that demonstrated the color temperature levels that produce better color resolution and rendering both to eye and camera.) Sixth, he collaborated on a study dealing with mercury containment in fluorescent lighting. The seventh and eight were consultations on color rendering in still photography and reflectivity of materials in the cargo bay of the Orbiter. The ninth problem that he worked on during the period of observation dealt with the design of an EVA helmet with LCD projection on its viewing surface.

The observations of consultative collaboration and advisory meetings led us to conclude that the expertise was indeed genuine. This conclusion was confirmed in interviews with a number of engineers and scientists at JSC who occasionally consult with CW. The extent of the expertise, however, was unexpected. Expert system technology does not work well with broad domains of knowledge and the issue of project suitability required that the project be narrowed to aspects of the domain of spacecraft lighting design. The domain represents one of CW's primary area of expertise. The design area may be segmented so that realizable goals may be set for pre-prototype problem evaluation and the initial steps of prototype construction.

### System Conceptualization

The specification of system requirements and the description of user populations are part of the system conceptualization phase of prototype development. Several two hour long knowledge engineering sessions were conducted with the expert. These were followed by interviews with potential users of a lighting design expert system. The combination of literature review and interviews lead to the formulation of the first cut in the system design: IVA (Intravehicular Activity) and EVA (Extravehicular Activity) lighting design. Although the lighting principles may be the same for the two, the heuristics, rules, and tests required appear to be quite different. IVA was selected. The next cut was made in accordance with major areas of spacecraft activity: general circulation, habitation, and workstations. Emergency and portable lighting was added as a fourth category of lighting design. Habitation was subdivided into crew quarters, ward room, waste management, personal hygiene, and showers. Workstation was divided into exercisers, general maintenance, health care, galley, windows, and science labs. The development of a prototype is typically focused on a restricted subdomain so that the potential of a full scale expert system can be demonstrated. CRT and other display and control workstation components were not included in an overall menu design at this time because instrument panel lighting entails unique design principles that add considerable additional complexity. The problem was deferred for future analysis. Also, an enumeration of the full set of components of lighting design to be described in a prototype and later incorporated in a final system was not attempted.

The conceptualization phase of development also entails the identification of representational and problem-solving methodologies that are suitable for the task. These issues will be briefly discussed below.

The major conclusions of the system conceptualization and development analysis are:

1. The development of a Spacecraft Lighting Advisor is a major expert system construction task and accordingly requires the methodologies and effort that is typical in large-scale knowledge based system development. The design experience and advice appropriate for building small knowledge systems does not apply and can be seriously misleading for the task at hand.

2. The development of a full prototype system will require approximately one man-year provided that the one or more members of development staff have the necessary training or experience in artificial intelligence programming and in the psychology of knowledge engineering.

3. The finished system may take an additional 1.5 man-years or longer depending on the number of features included.

#### Tool Selection

The most expensive expert system development packages, in general, tend to be the most powerful and possess the widest array of representational technologies for development. They also include procedures for effective execution, as well as the maintenance and modification of the knowledge base. Unfortunately, the most expensive packages also require the longest period to master, and full mastery only occurs when the tool is used in a real world project. There is also the problem of matching the package to the problem. For example, Clancey (1985) maintains that "when presented with a given 'knowledge engineering tool' such as EMYCIN (van Melle, 1979), we are still hard-pressed to say what kinds of problems it can be used to solve."

Typically, a large-scale problem, such as the one under consideration, will require hybrid knowledge representation including rules, frames and other object-oriented devices. At this stage of the evolution of the technology, however, no one can definitively select a knowledge engineering tool as appropriate for a given type of problem, let alone the one that is optimum for that problem. Yet, effective tool selection remains a process that can place boundaries around a project and render it computationally and financially tractable.

Early in the project, a decision was made against allocating any of the ten week project period to an in-depth evaluation of tools, as the required analysis can consume a project of short duration leaving little time for anything else. Consequently, an expert system shell developed for an

advanced graduate course in knowledge-based systems (Lachman, 1989) was modified for the task. The first menu of the shell contains a selection that jumps the program to a module patched into the main program. The patched module was used to demonstrate a partial prototype for spacecraft lighting. Fig. 3 shows a menu nested at the second level and contained in the patched code. The knowledge base developed for the pre-prototype system can be readily ported to a commercial heuristic programming environment once a tool has been selected that was designed for building large-scale systems. The final programming environment will require demonstrable advantages for maintenance, explanation facility, control of reasoning, and knowledge base representation of the final product.



Lyndon B. Johnson Space Center

<b>SPACECRAFT LIGHTING ADVISOR</b>		MAN-SYSTEMS DIVISION	
		ROY LACHMAN	SP-34

**SPACECRAFT LIGHTING  
ADVISOR**

Expert System: Select Area For Lighting Design

<p><b>GENERAL CIRCULATION</b></p> <p style="padding-left: 20px;">A PASSAGEWAYS</p> <p><b>HABITATION</b></p> <p style="padding-left: 20px;">B CREW QUARTERS C WARD ROOM D WASTE MANAGEMENT E PERSONAL HYGIENE F SHOWER</p>	<p><b>EMERGENCY AND PORTABLE</b></p> <p style="padding-left: 20px;">G GENERAL</p> <p><b>WORKSTATION</b></p> <p style="padding-left: 20px;">H EXERCISERS I GENERAL MAINTENANCE J HEALTH CARE K GALLEY L WINDOWS M SCIENCE LABS</p>
---	---

During Any Procedure: Press <F1> for Main Menu, < H > Exit to DOS

Figure 3.- A menu from the pre-prototype system.

The pre-prototype software includes elements of a backward chaining module that, when completed, would evaluate the design problem and interactively determine if the problem was appropriate for an expert system lighting advisor. This is not a trivial part of the project since design engineering can involve creative processes that cannot be implemented in automata at this time.

The overall plan for the expert system design is to try to duplicate processes generally used by human designers. The general theoretical orientation is the Newell and Simon (1972) information processing theory of problem solving. The problem space, in their theory, consists of a mental representation of the initial problem state, mental operators that change the problem state, and a termination criterion that is an acceptable solution to the problem at hand. The problem space consists of all the states or situations that, in principle, can be produced from the initial state by application of cognitive operators. A solution is any sequence of individual state-operator pairs that leads to a termination state. A problem-solving strategy is a method of search that identifies specific paths through the search space as worthy of consideration. Engineering design can be generally characterized in the search-space formulation. An expert system tries to capture from a human domain expert actual search-space strategies employed. A prototype should demonstrate the efficacy of the heuristic reduction of the search space and selective search operators that were copied from the human expert. It also should demonstrate the potential value of the final product to the sponsoring organization.

A pre-prototype system, in contrast, displays features in a pre-computational fashion that might be included in the final system. The goal is to demonstrate the logical adequacy of a given set of properties and a given approach. In addition, the pre-prototype may include a computational implementation of one or two of the features. Major properties of engineering design should be captured in the software. Engineers, typically, generate and test design alternatives for the search space of a given design specification. A rough design is produced and backtracking through the search space then produces incremental improvements in the design.

The simplest situation of problem solving in engineering design occurs when a previous design is implemented without modification; specifications are tested and a single unidirectional path through the search space is produced. Because of the project's time constraints, that type of solution was selected for the computational demonstration along with an uncomplicated area of lighting design on a spacecraft, the crew quarters. A sample

07-30-1988 16:16:19

VER=S14 RULE-BASE=SLA2.RB

Try To establish one of the following hypotheses from the rule base:

- 1 =NO-DESIGN-CHANGE:USE-STANDARD-15WATT-LUMINAIRE:TWO-PARALLEL-FIXTURES
- 2 =ADAPTIVE-DESIGN:CHANGE-LOCATION-OF-FIXTURES
- 3 =VARIANT-DESIGN:CHANGE-NUMBER-OF-FIXTURES-AND-LOCATION
- 4 =AN-ORIGINAL-DESIGN-REQUIRED:REDESIGN-LUMINAIRE

ENTER HYPOTHESIS NO. --->1

The System is trying to prove goal NO-DESIGN-CHANGE:USE-STANDARD-15WATT-LUMINAIRE:TWO-PARALLEL-FIXTURES

The System is trying to prove goal STANDARD-CREW-QUART.DIMENSIONS

The System is trying to prove goal STANDARD-D1

The System is trying to prove goal DIMENSION1(HEIGHT)=80IN.

All relevant rules have fired, can't prove DIMENSION1(HEIGHT)=80IN., which is not in STM-DB nor the consequent of a Rule.

IS THIS TRUE: DIMENSION1(HEIGHT)=80IN. <Y:N:D:W> --->Y

RULE 5 Deduces STANDARD-D1

Short-term memory (DATA BASE) now contains:

DIMENSION1(HEIGHT)=80IN. STANDARD-D1

The System is trying to prove goal STANDARD-D2

The System is trying to prove goal DIMENSION2(WIDTH)=41.5IN.

All relevant rules have fired, can't prove DIMENSION2(WIDTH)=41.5IN.

IS THIS TRUE: DIMENSION2(WIDTH)=41.5IN. <Y:N:D:W> --->Y

RULE 6 Deduces STANDARD-D2

Short-term memory (DATA BASE) now contains:

DIMENSION1(HEIGHT)=80IN. STANDARD-D1 DIMENSION2(WIDTH)=41.5IN. STANDARD-D2

The System is trying to prove goal STANDARD-D3

The System is trying to prove goal DIMENSION3(DEPTH)=39.5IN.

All relevant rules have fired, can't prove DIMENSION3(DEPTH)=39.5IN.

IS THIS TRUE: DIMENSION3(DEPTH)=39.5IN. <Y:N:D:W> --->Y

Figure 4. - A sample run of the pre-prototype expert system.

ORIGINAL PAGE IS  
OF POOR QUALITY

~~ORIGINAL PAGE~~  
~~COLOR PHOTO~~

IS THIS TRUE: STANDARD-WEIGHT-CONSTRAINTS-APPLY <Y:N:D:W> --->Y

The System is trying to prove goal STANDARD-HEAT-CONSTRAINTS-APPLY  
All relevant rules have fired, can't prove STANDARD-HEAT-CONSTRAINTS-APPLY,  
which is not in STM-DB nor the consequent of a Rule.

IS THIS TRUE: STANDARD-HEAT-CONSTRAINTS-APPLY <Y:N:D:W> --->Y

The System is trying to prove goal STANDARD-POWER-CONSTRAINTS-APPLY  
All relevant rules have fired, can't prove STANDARD-POWER-CONSTRAINTS-APPLY

IS THIS TRUE: STANDARD-POWER-CONSTRAINTS-APPLY <Y:N:D:W> --->Y

RULE 9 Deduces STANDARD-WEIGHT:HEAT:POWER-CONSTRAINTS

Short-term memory (DATA BASE) now contains:

DIMENSION1(HEIGHT)=80IN. STANDARD-D1 DIMENSION2(WIDTH)=41.5IN. STANDARD-D2  
DIMENSION3(DEPTH)=39.5IN. STANDARD-D3 STANDARD-CREW-QUART.DIMENSIONS  
STANDARD-WEIGHT-CONSTRAINTS-APPLY STANDARD-HEAT-CONSTRAINTS-APPLY STANDARD-  
POWER-CONSTRAINTS-APPLY STANDARD-WEIGHT:HEAT:POWER-CONSTRAINTS

RULE 1 Deduces NO-DESIGN-CHANGE:USE-STANDARD-15WATT-LUMINAIRE:TWO-PARALLEL-  
FIXTURES

Short-term memory (DATA BASE) now contains:

DIMENSION1(HEIGHT)=80IN. STANDARD-D1 DIMENSION2(WIDTH)=41.5IN. STANDARD-D2  
DIMENSION3(DEPTH)=39.5IN. STANDARD-D3 STANDARD-CREW-QUART.DIMENSIONS  
STANDARD-WEIGHT-CONSTRAINTS-APPLY STANDARD-HEAT-CONSTRAINTS-APPLY STANDARD-  
POWER-CONSTRAINTS-APPLY STANDARD-WEIGHT:HEAT:POWER-CONSTRAINTS NO-DESIGN-  
CHANGE:USE-STANDARD-15WATT-LUMINAIRE:TWO-PARALLEL-FIXTURES

FINAL RESULT: NO-DESIGN-CHANGE:USE-STANDARD-15WATT-LUMINAIRE:TWO-PARALLEL-  
FIXTURES

07-30-1988 16:16:51 NORMAL TERMINATION

RULE 7 Deduces STANDARD-D3

Short-term memory (DATA BASE) now contains:

DIMENSION1(HEIGHT)=80IN. STANDARD-D1 DIMENSION2(WIDTH)=41.5IN. STANDARD-D2  
DIMENSION3(DEPTH)=39.5IN. STANDARD-D3

RULE 8 Deduces STANDARD-CREW-QUART.DIMENSIONS

Short-term memory (DATA BASE) now contains:

DIMENSION1(HEIGHT)=80IN. STANDARD-D1 DIMENSION2(WIDTH)=41.5IN. STANDARD-D2  
DIMENSION3(DEPTH)=39.5IN. STANDARD-D3 STANDARD-CREW-QUART.DIMENSIONS

The System is trying to prove goal STANDARD-WEIGHT:HEAT:POWER-CONSTRAINTS

The System is trying to prove goal STANDARD-WEIGHT-CONSTRAINTS-APPLY

All relevant rules have fired, can't prove STANDARD-WEIGHT-CONSTRAINTS-  
APPLY, which is not in STM-DB nor the consequent of a Rule.

Figure 4.(continued) - A sample run of the pre-prototype expert system.

demonstration run is shown in Fig 4.

It should be possible to develop a spacecraft lighting advisor in an incremental fashion starting with the least problematic design areas and working up to the most difficult. Successful development of a heuristic system that evaluates the difficulty and feasibility of design problems for solution by an expert system is not guaranteed at this stage in the evolution of the technology.

#### SUMMARY

In this paper, we have examined the problems associated with the loss of specialized expertise and institutional memory through the attrition of personnel. The costs to the NASA organization were examined and several possible solutions were critiqued. Two solutions were selected for detailed examination: hypertext and expert systems. A paradigmatic case was selected in the person of an in-house expert on spacecraft lighting and windows. Sample computational implementations were produced and described. These were determined to offer reasonable solutions to at least part of the problem.

## REFERENCES

- Clancey, W. J. (1985). Heuristic classification. Artificial Intelligence, 27, 289-350.
- Conklin, Jeff (1987). Hypertext: An Introduction and Survey. IEEE Computer, 17-40, (September).
- Kaufman, J. E. and Christensen, J. F. (1984). (Eds.) IES Lighting Handbook: Reference Volume. New York: Illumination Engineering Society of North America.
- Kaufman, J. E. and Christensen, J. F. (1987). (Eds.) IES Lighting Handbook: Application Volume. New York: Illumination Engineering Society of North America.
- Lachman, R. (1989). Expert Systems: A cognitive science approach. Behavior Research Methods, Instrumentation, and Computers, 21, in review.
- Newell, A. and Simon, H. A. (1972). Human problem solving. Englewood Cliffs, NJ: Prentice-Hall.
- Van Melle, W. A domain-independent production system for consultation programs. Proceedings of the Sixth International Joint Conference on Artificial Intelligence. Pages 815-823, August, 1981.
- Waterman, D.A. (1986) A Guide to Expert Systems. Reading, MA: Addison-Wesley.

N89 - 20063

DEVELOPMENT OF PARALLEL ALGORITHMS FOR ELECTRICAL POWER  
MANAGEMENT IN SPACE APPLICATIONS

Final Report

NASA/ASEE Summer Faculty Fellowship Program--1988

Johnson Space Center

Prepared By:	Frederick C. Berry
Academic Rank:	Assistant Professor
University & Department:	Louisiana Tech University Dept. of Electrical Engineering Ruston, Louisiana 71291

NASA/JSC

Directorate:	Engineering
Division:	Avionics Systems
Branch:	Avionics Integration
JSC Colleague:	Tom Jeffcoat
Date Submitted:	August 12, 1988
Contract Number:	NGT 44-005-803

## ABSTRACT

This paper is concerned with the application of parallel techniques for electrical power system analysis. The Newton-Raphson method of load flow analysis was used along with the decomposition-coordination technique to perform load flow analysis. The decomposition-coordination technique enables tasks to be performed in parallel by partitioning the electrical power system into independent local problems. Each independent local problem represents a portion of the total electrical power system on which a load flow analysis can be performed. The load flow analysis is performed on these partitioned elements by using the Newton-Raphson load flow method. These independent local problems will produce results for voltage and power which can then be passed to the coordinator portion of the solution procedure. The coordinator problem uses the results of the local problems to determine if any correction is needed on the local problems. The coordinator problem is also solved by an iterative method much like the local problem. The iterative method for the coordination problem will also be the Newton-Raphson method. Therefore, each iteration at the coordination level will result in new values for the local problems. The local problems will have to be solved again along with the coordinator problem until some convergence conditions are met.

## INTRODUCTION

This paper will use the decomposition-coordination technique which enables task to be performed in parallel when solving sets of nonlinear equations(1, 4, 5, 6). Sets of nonlinear equations occur when performing a load flow analysis on an electrical power system(3, 7, 8). Load flow is the solution of an electrical network that gives the values of currents, voltages, and power flows at every bus (node) in the electrical power system(3, 7, 8). In the load flow problem nonlinear relationships between voltage and power occur at each bus(3, 7, 8). The values of the voltage and power must be solved for at each bus so that the response of the electrical power system can be determined. With the increase in the size of new proposed space based electrical power systems, it will become necessary to have very fast simulation (solution) of these systems. Very fast simulation of the electrical power system will aid in the evaluation of the system performance by decreasing the speed of calculations(1, 3, 4, 6, 7, 8).

## LOAD FLOW ANALYSIS

The Newton-Raphson method for performing the load flow calculation was used(2, 3, 7, 8). Taylor series expansion for a function of two or more variables is the basis of the Newton-Raphson method. Partial derivatives of order greater than 1 are neglected in the series terms of the Taylor series expansion. The Newton-Raphson method was used because it calculates corrections while taking into account all other interactions. The number of iterations required by the Newton-Raphson method using bus admittances is practically independent of the number of buses(3, 7, 8). For these reasons shorter computer time for a solution of the load flow problem could occur when analyzing large electrical power systems(3, 7, 8).

The solution of the load flow problem is initiated by assuming voltage values for all buses except the slack bus(3, 7, 8). The slack bus is the point at which the voltage is specified and remains fixed. The voltage at the slack bus is fixed because the net power flow of the system cannot be fixed in advance until the load flow study is complete(3, 7, 8). The power calculation at the slack bus supplies the difference between the specified real power

ORIGINAL PAGE IS  
OF POOR QUALITY

into the system at the other buses and the total system output plus losses(3, 7, 8). The Newton-Raphson method for load flow analysis will be used to solve the load flow problem at the local and coordinator problem levels(2, 3, 7, 8).

## DECOMPOSITION-COORDINATION METHOD

The decomposition-coordination method enables tasks to be preformed in parallel by partitioning the electrical power system into independent local problems(1, 4, 6). The power system presented in Figure 1 was solved by defining three local problems and a coordinator problem(1, 4, 6).

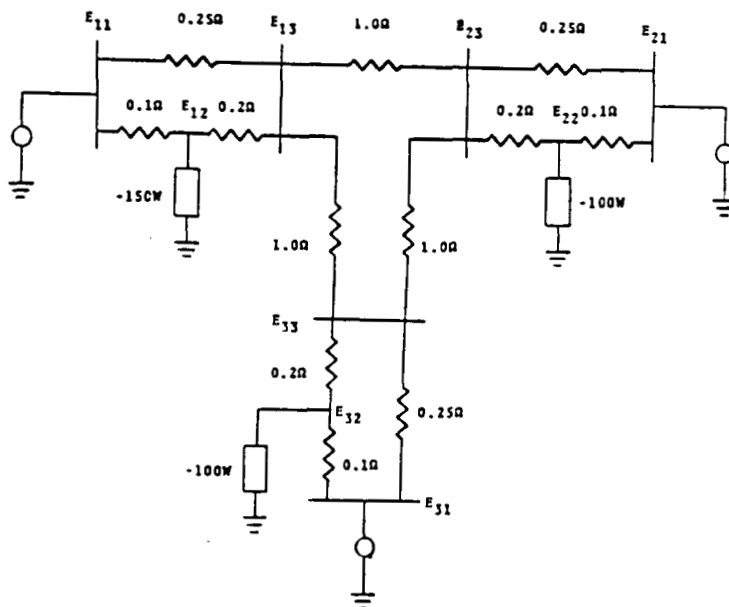


FIGURE 1.-CIRCUIT DIAGRAM OF POWER SYSTEM TO BE ANALYZED.

The local problems were solved by the Newton-Raphson procedure (1, 2, 3, 4, 6, 7, 8):

$$v_i^{n+1} = v_i^n - \left[ \frac{\partial F_i(p_i, v_i)}{\partial v_i} \Big|_{v_i=v_i^n} \right]^{-1} F_i(p_i^n, v_i^n) \quad . (1)$$

Starting with an initial guess for the voltage, equation (1) was used successively until convergence was achieved(1, 2, 3, 4, 6, 7, 8). Convergence is achieved when the power equations are in balance at each bus(3, 7, 8). This is represented by the following equation:

$$F_i (p_i, v_i) = 0 \quad .(2)$$

The system of equations which describe the three local problems are given in Table 1.

TABLE 1.-LOCAL PROBLEM EQUATIONS.

$$F(P_{11}, V_{11}) = -P_{11} + 14E_{11}^2 - 10E_{11}E_{12} - 4E_{11}E_{13} = 0$$

$$F(P_{12}, V_{12}) = -P_{12} - 10E_{12}E_{11} + 15E_{12}^2 - 5E_{12}E_{13} = 0$$

$$F(P_{13}, V_{13}) = -P_{13} - 4E_{13}E_{11} - 5E_{13}E_{12} + 9E_{13}^2 = 0$$

$$F(P_{13}, V_{13}) = -P_{13} + 2E_{13}^2 - E_{13}E_{23} - E_{13}E_{33} = 0$$

$$F(P_{23}, V_{23}) = -P_{23} - E_{23}E_{12} + 2E_{23}^2 - E_{23}E_{33} = 0$$

$$F(P_{33}, V_{33}) = -P_{33} - E_{33}E_{13} - E_{33}E_{23} + 2E_{33}^2 = 0$$

$$F(P_{21}, V_{21}) = -P_{21} + 14E_{21}^2 - 10E_{21}E_{22} - 4E_{21}E_{23} = 0$$

$$F(P_{22}, V_{22}) = -P_{22} - 10E_{22}E_{21} + 15E_{22}^2 - 5E_{22}E_{23} = 0$$

$$F(P_{23}, V_{23}) = -P_{23} - 4E_{23}E_{21} - 5E_{23}E_{22} + 9E_{23}^2 = 0$$

The solution for the local problems are usually not available in an explicit form. Therefore, the coordinator problem must be solved iteratively. This implies that for each iteration at the coordinator level, new values for the input to the local problem will result(1, 4, 6).

The Newton-Raphson procedure was also used for the solution of the coordinator problem:

$$v_k^{m+1} = v_k^m - \left[ \frac{\partial F_k(p_k, v_k)}{\partial v_k} \bigg|_{v_k=v_k^m} \right]^{-1} F_k(p_k^m, v_k^m) \quad .(3)$$

Starting with an initial guess for the voltage (these values should come from the last solution of the local problem), equation (3) was used successively until convergence was achieved (1, 2, 3, 4, 6, 7, 8). Convergence is achieved just like that of the local problem when the power equations balance to zero (or within a preset error). This is represented by the following equations:

$$F_k(p_k, v_k) = 0 \quad .(4)$$

The equations which describe the coordinator problem are given in Table 2.

TABLE 2.-COORDINATOR PROBLEM EQUATIONS.

$$F(P_{31}, V_{31}) = -P_{31} + 14E_{31}^2 - 10E_{31}E_{32} - 4E_{31}E_{33} = 0$$

$$F(P_{32}, V_{32}) = -P_{32} - 10E_{32}E_{31} + 15E_{32}^2 - 5E_{32}E_{33} = 0$$

$$F(P_{33}, V_{33}) = -P_{33} - 4E_{33}E_{31} - 5E_{33}E_{32} + 9E_{33}^2 = 0$$

#### EXAMPLE SOLUTION AND RESULTS

To solve the load flow problem of Figure 1 the local and coordinator problems of Table 1 and 2 were first solved by the following process:

1. There were three local problems designated and slack buses were assigned (given Table 3).

ORIGINAL PAGE IS  
OF POOR QUALITY

2. The local problems were then solved to determine what value of voltage and power must be present at the buses of each local system to meet the load requirement of each local problem.
3. The coordinator problem was then solved by assigning a slack bus(given Table 3) and using the most recent values of the voltage and power from the local problem.

This process was continued until the local and coordinator problems converged on a value of voltage for all buses. The results of this process are given in Table 3.

TABLE 3.-SOLUTION OF THE LOAD FLOW PROBLEM FOR FIGURE 1.

Local Problem Iteration 1.					
Bus	Power	Voltage	Bus	Power	Voltage
11	120.000	30.588	13	-30.579	30.303
12	-150.000	30.228	23	20.017	30.595
13	30.579	30.500*	23	-20.017	30.500*

Coordinator Problem Iteration 1.	
Bus	Power
31	112.051
32	-100.000
33	-11.213

Local Problem Iteration 2.					
Bus	Power	Voltage	Bus	Power	Voltage
11	120.000	30.388	13	-31.703	29.797
12	-150.000	30.026	23	18.874	30.359
13	31.703	30.303*	23	-18.874	30.595*

Coordinator Problem Iteration 2.	
Bus	Power
31	114.446
32	-100.000
33	-13.399

Local Problem Iteration 3.					
Bus	Power	Voltage	Bus	Power	Voltage
11	120.000	29.882	13	-31.763	29.785
12	-150.000	29.515	23	18.877	30.348
13	31.763	29.797*	23	-18.877	30.299*

Coordinator Problem Iteration 3.	
Bus	Power
31	114.507
32	-100.000
33	-13.458

Local Problem Iteration 4.					
Bus	Power	Voltage	Bus	Power	Voltage
11	120.000	29.871	13	-31.765	29.785
12	-150.000	29.503	23	18.880	30.348
13	31.765	29.785*	23	-18.889	30.348*

Coordinator Problem Iteration 4.	
Bus	Power
31	114.498
32	-100.000
33	-13.449

\*Slack Bus

As can be seen from the results of Table 3 convergence of the values of voltage and power have occurred to within an error of 0.01. To continue this solution procedure for 1 more iteration will result in convergence at errors of less than 0.001 for the voltage and power.

### CONCLUSION

This research used the decomposition-coordination technique along with the Newton-Raphson method to perform load flow analysis. It has been shown that parallel techniques can be used to solve standard power system problems. The solution procedure required the formulation of a set of local problems which could be run in parallel. The results of these local problems were then passed to the coordinator problem to determine if any correction is needed to the local problem. This process was continued until convergence conditions were met.

## REFERENCES

1. Buro, R., 1986, Application Of Parallel Processing To RELAP, Parallel Processing Techniques For Simulation, edited by M. G. Singh, A. K. Allinda, and B. K. Daniels.
2. Carnahan, B., Luther, H. A., and Wilkes, J. O., 1969, Applied Numerical Methods, New York, John Wiley and Sons, Inc.
3. Glover, J. D. and Sarma, M., 1987, Power System Analysis And Design, Boston, PWS Publishers.
4. Malinowskt, K., Allidina, A. Y., and Singh, M. G., 1986, Decomposition-Coordination Techniques For Parallel Simulation, Parallel Processing Techniques For Simulation, edited by M. G. Singh, A. Y. Allidna, and B. K. Daniels.
5. Perrott, R. H., 1987, Parallel Programming, Wokingham, England, Addison-Wesley Publishing Company.
6. Rice, John R., 1987, Parallel Methods For Partial Differential Equations, The Characteristics Of Parallel Algorithms, edited by Lean H. Jamieson, Dennis B. Gannon, and Robert J. Douglass.
7. Stevenson, W. D. Jr., 1975, Elements Of Power Systems Analysis, New York, McGraw-Hill.
8. Wood, A. L. and Bruce, W. F., 1984. Power Generation Operation And Control, New York, John Wiley and Sons.

AUTOMATIC VOICE RECOGNITION USING  
TRADITIONAL AND ARTIFICIAL  
NEURAL NETWORK APPROACHES

Final Report

NASA/ASEE Summer Faculty Fellowship Program-1988

Johnson Space Center

Prepared By:	Nazeih M. Botros, Ph.D.
Academic Rank:	Assistant Professor
University & Department:	Southern Illinois University Department of Electrical Engineering Carbondale, Illinois 62901

NASA / JSC

Directorate:	Engineering
Division:	Tracking and Communications
JSC Colleague:	William Jordan
Date Submitted:	August 19, 1988
Contract Number:	NGT 44-005-803

~~PREVIOUS PAGE BLANK - NOT PRINTED~~

## ABSTRACT

The main objective of this research is to develop an algorithm for isolated-word recognition. This research is focused on digital signal analysis rather than linguistic analysis of speech. Features extraction is carried out by applying a Linear Predictive Coding (LPC) algorithm with order of 10. Continuous-word and speaker independent recognition will be considered in future study after accomplishing this isolated word research

To implement and test the proposed algorithm a microcomputer-based data acquisition system has been designed and constructed. The system digitizes the voice signal, after passing through a 100 c/s-3.8 Kc/s band pass filter, with a sample rate of 8 KHz and stores the digitized data into a 64Kx10-dynamic random access memory (DRAM) buffer. A squelch circuit consists mainly of comparators (741s85) detects the beginning of the spoken word. The end of the word is detected by a software algorithm based on comparing the speech energy with a precalculated threshold. A flag signals the end of sampling and the data is transferred from the buffer to an IBM-PC where it is segmented into frames each, 30 millisecond long, and the LPC coefficients are calculated.

To examine the similarity between the reference and the training sets, two approaches are explored. The first is implementing traditional pattern recognition techniques where a dynamic time warping algorithm is applied to align the two sets and calculate the probability of matching by measuring the Euclidean distance between the two sets. The second is implementing a backpropagation artificial neural net model with three layers as the pattern classifier. The adaptation rule implemented in this network is the generalized least mean square (LMS) rule.

The first approach has been accomplished. A vocabulary of 50 words was selected and tested, the accuracy of the algorithm was found to be around 85%. The second approach is in progress at the present time. The topology of the backpropagation model consists of three layers: input, hidden, and output. The actual output of each node is calculated using a sigmoid nonlinearity function of the inner product of the weight and the input; the weights are adapted by using the formula  $W_{ij}^{new} = W_{ij}^{old} + u E_j X_i$  where  $u$  is the gain factor,  $E_j$  is the error, and  $X_i$  is the input. The network is being simulated on a PC.

## INTRODUCTION

For more than a decade the United States Government, foreign countries especially Japan, private corporations, and universities have been engaged in extensive research on human-machine interaction by voice. The benefits of this interaction is especially noteworthy in situations when the individual is engaged in such hands/eyes-busy task, or in low light or darkness, or when tactile contact is impractical/impossible. These benefits make voice control a very effective tool for space-related tasks. Some of the voice control applications that have been studied in NASA-JSC are: VCS Flight experiments, payload bay cameras, EVA heads up display, mission control center display units, and voice command robot. A special benefit of voice control is in zero gravity condition where voice is a very suitable tool in controlling space vehicle equipment.

Automatic speech recognition is carried out mostly by extracting features from the speech signal and storing them in reference templates in the computer. These features carry the signature of the speech signal. These reference templates contain the features of a phoneme, word, or a sentence, depending on the structure of the recognizer. If a voice interaction with the computer takes place, the computer extracts features from this voice signal and compares it with the reference templates. If a match is found, the computer executes a programmable task such as moving the camera up or down.

Several digital signal processing algorithms are available for speech feature extraction. The efficiency of the current algorithms is limited by: hardware restriction, execution time, and easiness of use. Some of these algorithms are: Linear Predictive Coding (LPC), Short-time Fourier Analysis, and Cepstrum analysis. Among these algorithms, LPC is the most widely used since it is easy to use, has short execution time, and do not require large memory storage. However, this algorithm has several limitations due to the assumptions upon which it is based upon.

Current speech recognition technology is not sufficiently advanced to achieve high performance on continuous spoken input with large vocabularies and/or arbitrary speakers. A major obstacle in achieving such high performance is the limited capability of the traditional pattern recognition (classifier) algorithms that are currently implemented. Due to this limited capability, and considering the fact that humans have a fascinating capability of recognizing the spoken words, researchers have started to explore the

possibility of implementing human-like models, or what is known as artificial neural networks, as the pattern classifiers.

Neural net models have the greatest potential in area such as speech and image recognition where many hypotheses are pursued in parallel, high computation rates are required, ability to learn is desired, and the current best systems are far from equaling human performance. Most neural net algorithms adapt connection weights in time to improve performance based on current results. Adaptation or learning is a major focus of neural net research. The ability to adapt and continue learning is essential in area such as speech recognition where training data is limited and new talkers, new words, new dialects, new phrases, and new environments are continuously encountered.

### LINEAR PREDICTIVE CODING (LPC)

Feature extraction in our research is carried out through a Linear Predictive Coding algorithm. The signal corresponding to a spoken word is segmented into frames each 30 milliseconds long and the algorithm is applied to replace each frame with 10 coefficients. In the following, we briefly review the LPC algorithm. Details of this algorithm can be found elsewhere [1-9]. This algorithm is built on the fact that there is a high correlation between adjacent samples of the speech signal in the time domain. This fact means that an nth sample of speech signal can be predicted from previous samples. The correlation can be put in a linear relationship as:

$$\hat{Y}_n = a_1 Y_{n-1} + a_2 Y_{n-2} + \dots + a_p Y_{n-p} \quad \dots\dots\dots(1)$$

where p is the order of analysis. Usually p ranges from 8 to 12.  $\hat{Y}_n$  is the predicted value of speech at time n and a's are the linear predictive coefficients. The prediction error  $E_n$  that resulted from the above linear relationship is:

$$E_n = \sum_{i=0}^p a_i Y_{n-i}, \quad a_0 = 1. \quad \dots\dots\dots(2)$$

To find the predictive coefficients which give least mean square error, the above equation is squared, partially differentiated with respect to a's, and time average term by term. The result is p equations in p unknowns as shown below:

$$\begin{bmatrix} r_0 & r_1 & r_2 & \dots & r_{p-1} \\ r_1 & r_0 & r_1 & \dots & r_{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{p-1} & r_{p-2} & r_{p-3} & \dots & r_0 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_p \end{bmatrix} \dots\dots\dots(3)$$

with

$$\begin{aligned} r_0 &= \overline{Y_n Y_n}, \\ r_j &= \overline{Y_n Y_{n+j}} = \overline{Y_{n-j} Y_n} \end{aligned} \dots\dots\dots(4)$$

where  $r_j$  is a correlation coefficient of waveform  $\{y_n\}$  and  $r_{-j} = r_j$  by the assumptions of stationary state of  $y_n$ . The coefficients  $a_i$ 's exist only if the matrix in equation 3 is a positive definite. To ensure that this condition is satisfied,  $y_n$  is multiplexed by a time window  $W_n$ . This multiplexing makes  $y_n$  exist in a finite interval from 0 to  $N-1$ , where  $N$  is the interval of the Window; a stable solution for equation 3 is always obtained. Accordingly,  $r_j$  can be written as:

$$r_j = \frac{1}{N} \sum_{n=0}^{N-j-1} Y_n Y_{n+j} W_n W_{n+j} \dots\dots\dots(5)$$

In our study, a Hamming window is implemented. Calculation of the correlation coefficients by window multiplexing is called the correlation method.  $a_i$ 's correspond to the resonance frequencies of the signal, and if  $p$ , the order of the analysis is selected correctly, these  $a_i$ 's represent the formants, frequencies at which peaks of the power spectrum of the speech signal occur. A block diagram representing an algorithm for voice recognition based on LPC analysis is shown in Figure 1.

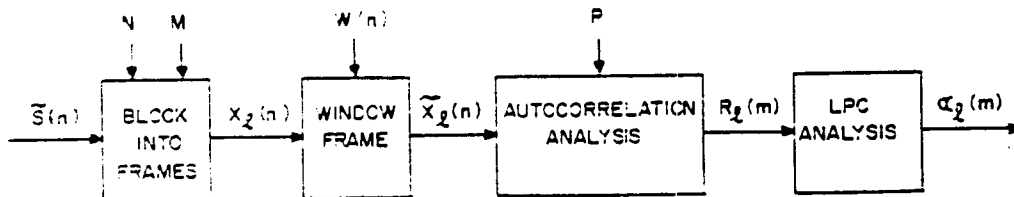


FIGURE 1. CALCULATION OF THE LPC COEFFICIENTS [5]

### EXPERIMENTAL DESIGN

To calculate the LPC coefficients and test the performance of the proposed algorithm, a microprocessor-based data acquisition system has been designed and constructed. Figure 2 shows a block diagram of the system. The system is

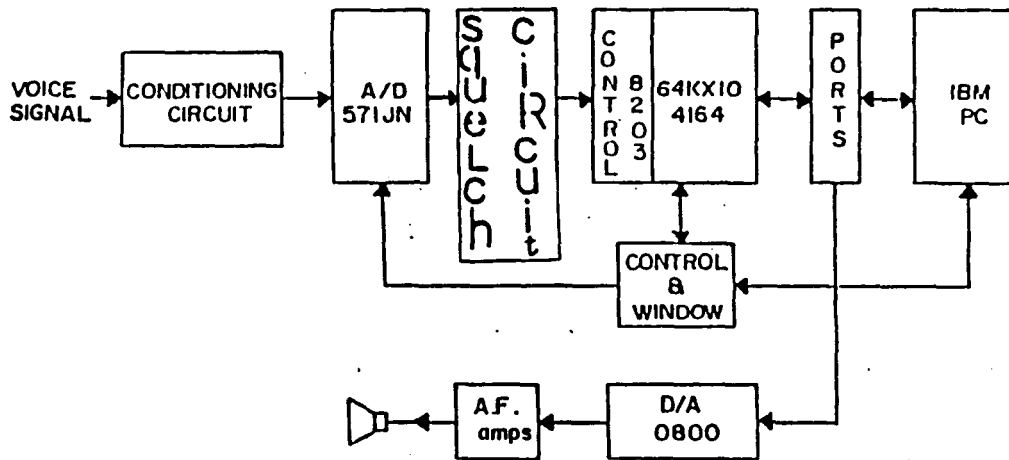


FIGURE 2. A BLOCK DIAGRAM OF THE DATA ACQUISITION SYSTEM.

designed to lay the foundation for further expansion and enhancement for more sophisticated microprocessor-based speech identification/ recognition research. The system receives the voice signal through a microphone coupled with an audio amplifier. The voice signal passes through an active multiple feedback bandpass filter with a 3db bandwidth of 3.5 KHz approximately. The output signal of the filter is applied to the analog-to-digital converter where it is digitized with a sampling rate of 8 KHz. A squelch circuit is constructed to detect the beginning of the utterance and accordingly activates a temporary storage buffer to store the digitized data. The buffer consists of DRAMS with maximum capacity of 64 Kbyte. A hardware flag (the output bit of a flip flop) signals the end of sampling and the data is transferred to the microprocessor where digital signal analysis and pattern recognition algorithms are applied. The digital-to-analog converter, power amplifier and loud speaker are used to verify the storage. If the output of this circuit matches the original signal, then the storage is successful. The system has been tested successfully by the aid of a function generator. Details of the hardware of the

system can be found elsewhere [9].

### PATTERN RECOGNITION

To recognize the spoken word, an algorithm that compares between the reference and the training patterns to see whether they match or not should be developed. Two approaches will be discussed here. The first is a traditional pattern classifier approach where training pattern is aligned in time with the reference pattern and the Euclidean distance between the two patterns is calculated and taken as the probability of matching. The second is based on implementing a backpropagation artificial neural network as the pattern classifier. In the following, we discuss briefly the two approaches.

#### A. Dynamic Time Warping (DTW)

The dynamic time warping (DTW) algorithm finds the "optimal" (least cost) warping path  $w(n)$  which minimizes the accumulated distance,  $D$ , between training and reference patterns, subject to a set of path and endpoint constraints. The dynamic programming algorithm is based upon the fact that the optimal path to point  $(i,j)$  in the two dimensional matrix illustrated in Figure 3 must pass through either the point  $(i-1,j)$ , or  $(i-1,j-1)$ , or  $(i,j-1)$ . The minimum accumulated distance to point  $(i,j)$  is then given by:

$$D(i,j) = \text{Dist}(i,j) + \text{Min}\{D(i-1,j), D(i-1,j-1), D(i,j-1)\} \quad \dots\dots\dots (11)$$

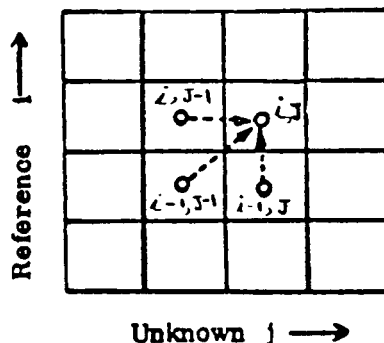


FIGURE 3. DYNAMIC TIME ALIGNMENT

where  $\text{Dist}(i,j)$  is the distance between the reference and training pattern at time  $j$ . The algorithm recursively computes this distance column by column to determine the minimum accumulated distance to the point  $(M,N)$ , where  $M$  is

the number of frames in the reference and  $N$  is the number of frames in the unknown (training). This path results in a time alignment in which the reference word has the maximum acoustic similarity with the input.

#### B. Artificial Neural Network [10-18]

Artificial neural network is a non-algorithmic information processing structure based on the architecture of our biological nervous system. The structure is composed of a massive number of processing elements operating in a predetermined parallel operation. The processing elements are connected by links with variable weight. The topology of each network determines the way each processor is connected to the other. The link can be excitatory, inhibitory, or has no effect on the activity of the processing element. See Figure 4. The primary processing at each element consists of

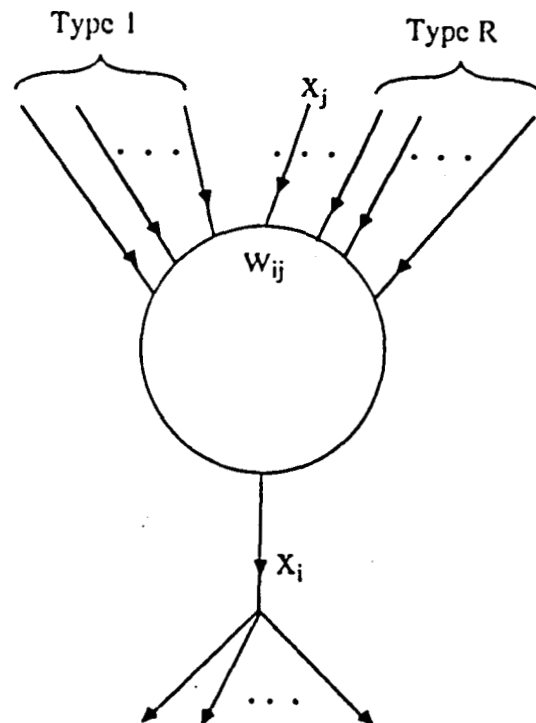


FIGURE 4. A PROCESSING ELEMENT.

the calculation of weighted sums of the form  $f(W_{ij}, X_j)$  and weight changes of the form  $W_{ij}^{new} = G(W_{ij}^{old}, X_i, X_j, \dots)$ . the function  $f$  is usually a nonlinear function,  $W_{ij}$  is the weight of the link from element  $i$  to element  $j$  and  $X_i$  is the input to element  $i$ . See Figure 5. Figure 6 shows some of the most popular neural

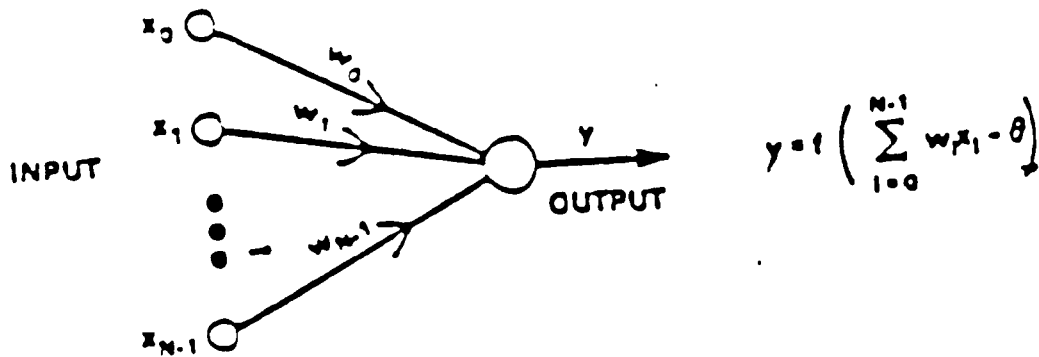


FIGURE 5. OUTPUT, INPUT, AND CONNECTION FOR A PROCESSING ELEMENT

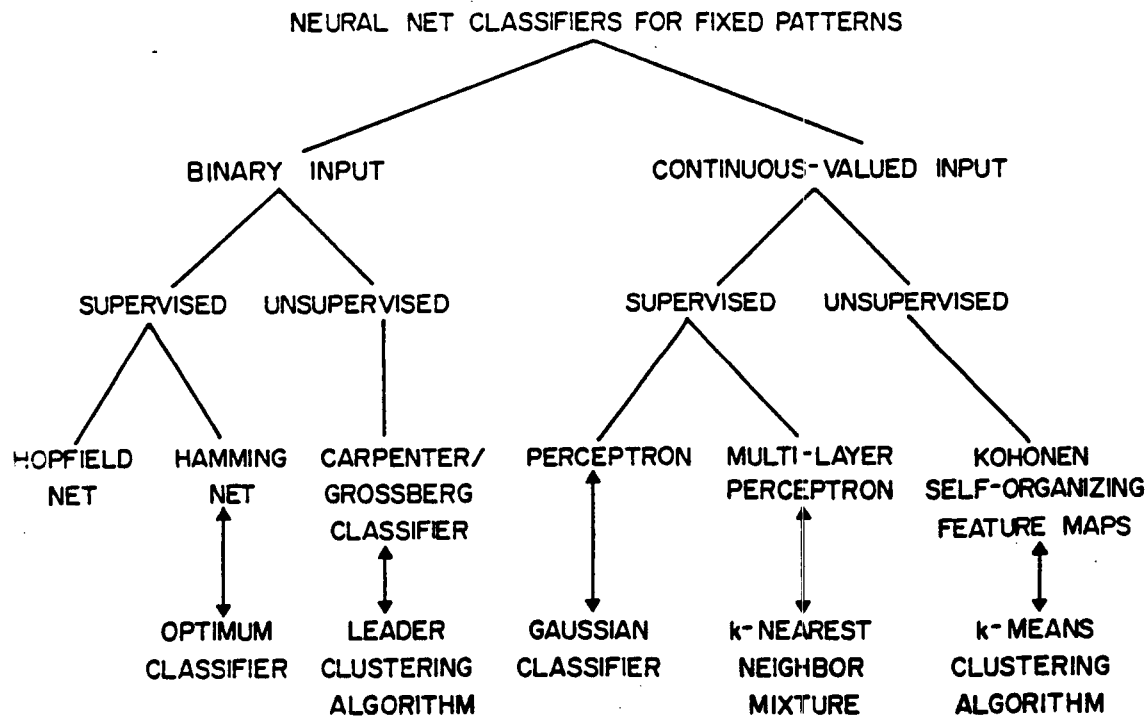
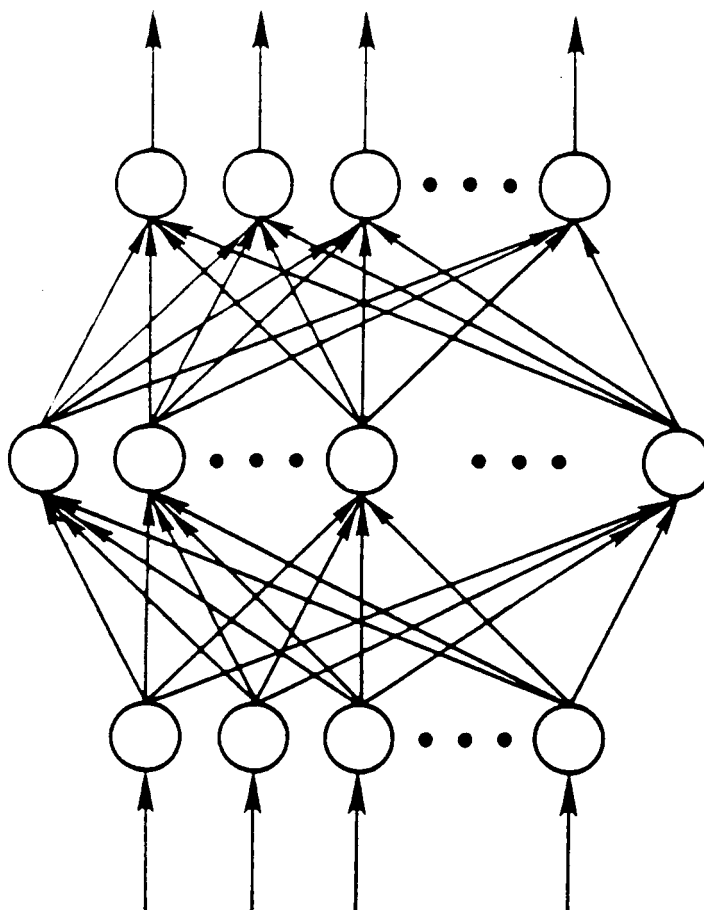


FIGURE 6. CLASSIFICATION OF ARTIFICIAL NEURAL NETWORKS [10]

net models that can be used as classifiers with the classical algorithms that are most similar to the neural net models are listed along the bottom[10]. As shown in this figure, these nets are first divided according to whether the input is binary or continuous-value. Second, they are divided according to whether they need supervision during training or not. Since a multi-layer perceptron backpropagation model is implemented in our study, we discuss in the following section the topology and the learning rule of that model.

#### The Backpropagation Model [10,14,15]

Figure 7 shows the topology of the backpropagation  
Output Patterns



Input Patterns

FIGURE 7. BACKPROPAGATION NETWORK.

model. The model consists of 3 layers (input, hidden, and output). each node in the input layer is connected to every node in the hidden layer also each node in the hidden layer is connected to every node in the output layer. The input of the model is a continuous valued vector  $x_0, x_1, \dots, x_{N-1}$  representing either the LPC coefficients or the frequencies of the Formants which are calculated from the coefficients. Investigations, mainly experimental, will be carried out to see which input is most suitable for word recognition. The actual output  $y_0, y_1, \dots, y_{M-1}$  is calculated as:

$y = f(\sum W_{ij}X_j - \theta)$  where  $\theta$  is a predetermined threshold and the function  $f$  is the sigmoid nonlinearity:

$$f(a) = \frac{1}{1 + e^{-(a-\theta)}}$$

Training of the network is carried out by setting the output of the model to the desired output vector  $d_0, d_1, \dots, d_{M-1}$ . All elements of the desired output vector are set to zero except for that corresponding to the current input training word which is set to 1. The weights are adjusted recursively from the output nodes to the hidden nodes by the formula:

$$W_{ij}^{new} = W_{ij}^{old} + u \delta_j X'_i$$

where  $u$  is the gain factor,  $\delta_j$  is the error, and  $X'_i$  is either the output of node  $i$  or is an input. If node  $j$  is an output node, then

$$\delta_j = y_j(1-y_j)(d_j-y_j),$$

where  $d_j$  is the desired output of node  $j$  and  $y_j$  is the actual output. If node  $j$  is an internal hidden node, then

$$\delta_j = x_j'(1-x_j') \sum_k \delta_k w_{jk},$$

where  $k$  is over all nodes in the layers above node  $j$ .

## REFERENCES

1. S. Saito, and K. Nakato, 1985, Fundamental of Speech Signal Processing, Shuzo Saito and Kazuo Nakato, Academic Press, Inc.
2. F. Fallside, and W. Woods, 1985, Computer Speech Processing, Prentice Hall International (U.K.) Ltd.
3. F. Itakura, 1975, "Minimum Prediction Residual Principle Applied to Speech Recognition", IEEE Trans. Acoust., Speech, Signal, Processing, Vol. ASSP-23, pp. 67-72.
4. L.R. Rabiner, 1978, "On Creating Reference Template for Speaker Independent Recognition of Isolated Word", IEEE Trans. Acoust., Speech, Signal, Processing, Vol. ASSP-26, pp. 34- 42.
5. L.R. Rabiner, and S.E. Levinson, 1981, "Isolated and Connected Word Recognition-Theory and Selected Application", IEEE Trans. on Communication, Vol. comm. 29, No. 5.
6. W.A. Lea, (1979) "Trends in Speech Recognition." Prentice-Hall, Inc., Englewood Cliffs, New Jersey.
7. N.M. Botros, "Digital Signal Processing Algorithms for Automatic Voice Recognition." Final report, NASA/ASEE Summer Faculty Fellowship Program-1987.
8. T. Parsons, "Voice and Speech Processing", McGraw-Hill Book Company, New York.
9. N. Botros, P. Hsu, and K. Xu, "A microprocessor-based system for voice identification", Proceedings of the ISMM International Symposium on Software and Hardware Applications of Microcomputers, pp. 52-56, Fort Collins, Co. 1987.
10. R. Lippmann, " An Introduction to Computing With Neural Nets," IEEE-ASSP Magazine, 4-22, April 1987.
11. J.J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," Proc. Natl. Acad. Sci. USA, Vol. 79, 2554-2558, April 1982.
12. T. Kohonen, K. Masisara and T. Saramaki, "Phonotopic Maps -Insightful Representation of Phonological Features for Speech Representation," Proceedings IEEE 7th Inter. Conf. on Pattern Recognition, Montreal, Canada, 1984.

13. R.F. Lyon and E.P. Loeb, "Isolated Digital Recognition Experiments with a Cochlear Model," in Proceedings International Conference on Acoustics Speech and Signal Processing, ICASSP-87, Dallas, Texas, April 1987.
14. D. E. Rumelhart, G.E. Hinton, and R.J. Williams, "Learning International Representations by Error Propagation" in D.E. Rumelhart & J.L. McClelland (Ed.), Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 1: Foundations. MIT Press (1986).
15. D.E. Rumelhart, and J.L. McClelland, Parallel Distributed Processing: Explorations in the Microstructure of Cognition, MIT Press (1986).
16. T. Sejnowski and C.R. Rosenberg, "NETtalk: A Parallel Network That Learns to Read Aloud," Johns Hopkins Univ. Technical Report JHU/ERCS-86/01, 1986.
17. D.W. Tank and J.J. Hopfield, "Simple 'Neural' Optimization Networks: An A/D Converter, Signal Decision Circuit, and a Linear Programming Circuit," IEEE Trans. Circuits Systems CAS-33, 533-541, 1986.
18. B. Widrow and S.D. Stearns, Adaptive Signal Processing, Prentice-Hall, New Jersey (1985).

DEVELOPMENT OF AN ATMOSPHERIC MONITORING PLAN  
FOR SPACE STATION

Final Report

NASA/ASEE Summer Faculty Fellowship Program--1988

Johnson Space Center

Prepared By: Dennis M. Casserly, Ph.D., CIH  
Academic Rank: Assistant Professor  
University & Department: University of Houston - Clear Lake  
School of Natural and Applied Sciences  
Division of Life Sciences  
Houston, Texas 77058

NASA/JSC

Directorate: Space and Life Sciences  
Division: Medical Sciences  
Branch: Biomedical Laboratories  
JSC Colleague: Dane M. Russo, Ph.D.  
Date Submitted: October 15, 1988  
Contract Number: NGT 44-005-803

## ABSTRACT

An environmental health monitoring plan for Space Station will ensure crew health during prolonged habitation. The Space Station, "Freedom" will operate for extended periods, 90+ days, without resupply. A regenerative, closed loop life support system will be utilized in order to minimize resupply logistics and costs. Overboard disposal of wastes and venting of gases to space will be minimal. All waste materials will be treated and recycled. The concentrated wastes will be stabilized and stored for ground disposal. The expected useful life of the station (decades) and the diversity of materials brought aboard for experimental or manufacturing purposes, increases the likelihood of cabin contamination. Processes by which cabin contamination can occur include: biological waste production, material off-gassing, process leakage, accidental containment breach, and accumulation due to poor removal efficiencies of the purification units.

An industrial hygiene approach was taken to rationalize monitoring needs and to identify the substances likely to be present, the amount, and their hazard. This requires a thorough knowledge of the onboard processes, their products and by-products. Many factors influence the monitoring requirements for Space Station: the enclosed space, the recirculation of supply air, the experiences of past missions, the unique experimental and manufacturing facilities, and the interfacing of other modules with the U.S. core modules. Monitor development and selection will be complicated due to the many technologies competing for the life support systems, and the number of experimental payloads under development, each having their own unique monitoring requirements.

Monitoring options include: on-line sensors for process control and determining efficacy of the life support regeneration and purification units; monitors for specific components and contaminants critical for life support; monitors for surrogate parameters representative of contaminant groups likely to be encountered; broad spectrum analyzers capable of identifying and quantifying nearly any contaminant for leak detection and remedial action; manual aboard sampling and analysis; and sample collection/preservation with ground based analysis. Other factors considered in the monitoring plan include: results of on ground system tests; consequences of contaminant detection failure; instrument parameters such as weight, volume, reliability, specificity, detection limit, and maintenance requirements; crew time and effort; expendables; and waste production.

## INDUSTRIAL HYGIENE PERSPECTIVE

It is standard practice in industry to monitor for those agents in the workplace that may occur at levels approaching or exceeding safe limits. Defining those limits and deciding on which agents to monitor, calls for a systematic and thorough approach, often employed in the field of industrial hygiene.

Industrial Hygiene is the science and art devoted to the anticipation, recognition, evaluation, and control of environmental factors and stresses arising in or from the workplace that may cause sickness, impair health and well being, discomfort or inefficiency among workers, their families or members of the community at large.<sup>1</sup>

Work hazards should be anticipated before they exist and action must be taken to prevent their occurrence, so no individual is placed at undue risk. Recognition of potential hazards requires familiarity with the process and work operations involved, the maintenance of an inventory of agents, a periodic review of the job activities, and the effectiveness of control measures.

Evaluation principally involves performing the monitoring to determine the exposures, comparing the results with standards, and communicating the judgement on the degree of hazard with the individuals affected and those in authority to take corrective action. Monitoring is a continuing program of observation and judgement. Reasons for monitoring are: to determine exposure levels; to determine effectiveness of control measures; to detect process changes; to investigate complaints; and to confirm compliance with standards.

In monitoring, part of the environment is sampled and the quality of the whole environment is inferred. A number of factors must be considered in order to take a representative sample. In industry the concern is worker exposure, so air samples are taken in the breathing zone of the worker, specified as twelve inches from the nose. Data from such personal samples are better correlated with exposures than any other type because of large concentration gradients and awkward operator/machine interfaces that can occur in industry. Sampling general workroom air is effective when the emission rates are uniform and there is good mixing. However, when relying on area sampling, documentation of its correlation with personal samples must be available. Sampling at the operation usually yields the highest concentration and worst case exposures. Rest areas should be occasionally monitored to confirm a clean zone of zero exposure.

All individuals exposed at or above the action level must be monitored. The action level is some fraction of the standard below which exposures must remain, usually one-half or one-fourth the standard. It is prudent to monitor nearby workers and complainers, as they may be sensitive to the agent, or an unidentified process leak may be present, or a control device may have failed.

Sample duration depends on the detection limit of the technique used for analysis, the standard, and the estimated air concentration. The sample duration should represent an identifiable period, consistent with the standard, the work shift and the process. Common sample durations are eight hours, for time weighted average standards and ten minutes for peak concentrations. Continuous monitoring is required if ceiling values are likely to be exceeded or when conditions immediately dangerous to life and health may occur.

Monitoring results are compared with standards to determine the severity of the exposure and a plan of action. Standards which must be met are the Permissible Exposure Limits (PEL's) published by the Occupational Safety and Health Administration (OSHA).<sup>2</sup> These are legal standards and represent the minimum effort required to maintain a safe and healthful workplace. Other standards or guidelines are recommended by professional organizations and consensus groups. Most widely used are the Threshold Limit Values (TLV's) of the American Conference of Governmental Industrial Hygienists (ACGIH).<sup>3</sup> Most PEL's were adopted from TLV's and most are eight hour time weighted average or ceiling values. For each standard there is a criteria document which presents data on which the standard is based.

Control can be achieved by reducing the emission source, interrupting the air path, and protecting the receiver. Specific control methods are: elimination or substitution with a less harmful agent; process change; enclosure of the process or source; isolation in time or distance; wet methods to reduce dust loadings and to scrub gases; local exhaust ventilation to remove the contaminant at its source; dilution ventilation; housekeeping; adequate maintenance; training and education of workers; area and personal monitoring; and use of personal protective equipment, mainly as backup for harmful or life threatening situations.

Elimination, substitution, and process change are ground based decisions that must be made well before flight. Strict flight requirements are being relaxed to allow the use of off the shelf items for Space Station. Isolation, wet methods, and dilution are not compatible with space vehicles. During space flight, control methods will rely heavily on enclosure, venting to the trace contaminant control system, housekeeping, and monitoring to assure a healthy cabin environment.

### Supplied Breathing Air

On ground, when we encounter situations with probabilities of oxygen deficiencies or conditions immediately dangerous to life and health, breathing air must be supplied.<sup>4</sup> It must be Grade D or better, with specifications as follows: oxygen 19-23%, carbon monoxide (CO)  $\leq 20$  ppm, carbon dioxide (CO<sub>2</sub>)  $\leq 1000$  ppm, condensed hydrocarbons  $\leq 5$  mg/M<sup>3</sup>, and the water content must be stated. The air source is ambient air, supplied by a compressor to a delivery system or cylinder for storage. Any one of these

can be a source of contamination. A frequently encountered cause of morbidity from contaminated air is carbon monoxide, from the incomplete oxidation of hydrocarbons of the compressor fuel, exhaust, and lubricants. Another source is from over heating of charcoal filters used for purifying the intake air. For this reason, continuous carbon monoxide and/or temperature monitors are required on air supply compressors. The condensed hydrocarbons can cause lipoid pneumonia and decrease the gas exchange membrane surface in the lung. They may also oxidize to CO and CO<sub>2</sub> within the storage cylinders. The water content is important: if excessive, it can cause regulator valves to clog or freeze, and promote cylinder corrosion; and if the air is too dry then irritation of mucous membranes, eyes, nose, and throat can occur. The oxygen (O<sub>2</sub>) content should be checked routinely to confirm adequate concentrations. Industrial users check every cylinder of air before use, since several fatal incidents have occurred due to low oxygen concentrations.

### Confined Spaces

In confined spaces, an area is enclosed or partially enclosed with poor ventilation and mixing with the outside air is limited. Before entry, the atmosphere should be tested to ensure the O<sub>2</sub> concentration is 19.5-21%. Toxic, flammable, and oxygen displacing gases and vapors should also be monitored. Common contaminants to check are hydrogen sulfide (H<sub>2</sub>S), CO, CO<sub>2</sub>, and methane (CH<sub>4</sub>), plus any other material that is likely to be present because of prior storage. If conditions immediately dangerous to life and health are probable, then continuous monitoring is required, since conditions within a confined space can change rapidly.<sup>5</sup>

### Nuclear Submarines

Submarines are enclosed environments with controlled atmospheres and are capable of remaining submerged for extended periods. Each submarine is equipped with a monitoring system designed to measure gaseous atmospheric constituents which are important in life support, called the Central Atmosphere Monitoring System (CAMS-I). It continuously analyzes air from the main fan room for eight substances, with a mass spectrometer (MS) and an infrared analyzer (IR). The mass spectrometer looks at specific m/e in the range of 2-300 amu. The average failure time was reported to be 3500 hours.<sup>6</sup> It is considered a monitor and not a trace contaminant detector, since sensitivity was restricted to meet stability and dependability requirements. CAMS-II is under development.<sup>7</sup> It will use a scanning mass spectrometer to measure 12 substances continuously from multiple sample ports located throughout the submarine. The MS will measure:

non-methane aliphatic hydrocarbons	0-100 ppm
aromatic hydrocarbons	0-10 ppm
benzene	0-10 ppm
carbon dioxide	0-3.3 %
hydrogen	0-5 %
nitrogen	0-80 %
oxygen	0-25 %
refrigerant R-11	0-1200 ppm
refrigerant R-12	0-1200 ppm
trichloroethylene	0-100 ppm
water vapor	0-4 %

Since the MS cannot resolve CO and N<sub>2</sub>, as both have m/e of 28, an IR measures for CO. CO is a major concern in submarines because of combustion processes, smoking, cooking, and smoldering of activated charcoal filters designed to remove odors, refrigerants, and hydrocarbons. Refrigerants themselves are of little concern, however their decomposition products from compressors or fire are corrosive, toxic, and will poison the catalytic oxidizer. Also on board is a paramagnetic O<sub>2</sub> analyzer and a Dwyer CO<sub>2</sub> analyzer. A portable photoionization detector is used for hydrocarbons, it is referred to as the "trace gas analyzer." Earlier CAMS used a gas chromatograph with a flame ionization detector as a hydrocarbon detector in which 100 ml samples were injected. It was not compatible with submarine duty.<sup>8</sup> An assortment of Draeger detector tubes are used for leak detection and for CAMS backup.

#### MONITORING ON SHUTTLE

The air revitalization system and many other systems on the Shuttle are different from those proposed for Space Station.<sup>9</sup> On the Shuttle, CO<sub>2</sub> is removed by LiOH canisters which are changed when the CO<sub>2</sub> concentration is 50 mm of Hg, or every 12 hours. Activated charcoal filters are used to remove hydrocarbons and odors. Carbon monoxide is oxidized to CO<sub>2</sub> in a low temperature catalytic oxidizer, located downstream from the humidity and thermal control unit. The activated charcoal filters and the low temperature catalytic oxidizer are capable of removing most contaminants that may occur. Fresh air is supplied from cryogenic liquid nitrogen (N<sub>2</sub>) and (O<sub>2</sub>) stored on board. Thus on the Shuttle, continuous monitoring is done for total pressure, temperature, humidity, O<sub>2</sub>, and CO<sub>2</sub>. Air samples are taken in evacuated bottles, and by active collection on various sorbent tubes, such as charcoal, tenax, and molecular sieve. Those samples are post-flight analyzed in ground based laboratories. The charcoal and LiOH air purification filters are also often analyzed for contaminants.

## MONITORING NEEDS

Experiences of past missions and ground based systems tests have identified a number of health concerns that should be addressed in a monitoring plan for Space Station. Paramount is the flight and post flight health complaints of the crews: headache; irritation of the eyes and upper respiratory tract; and odor complaints, symptomatic of noxious air.<sup>10</sup> Early missions had insufficient monitoring data for evaluation, which indicated a need for a more comprehensive monitoring system. Analyses of activated carbon and lithium hydroxide filters of the atmospheric revitalization systems, and the active sampling and analysis for air contaminants of later missions have identified over 250 contaminants in spacecraft air.<sup>11</sup> Most were observed at trace levels, well below the Spacecraft Maximum Allowable Concentration (SMAC). Others may have elicited symptoms among crew members, may accumulate to harmful levels, or may have potential to poison the spacecraft life support system.

Nitrogen tetroxide ( $N_2O_4$ ), hydrazine, and monomethyl hydrazine are the main liquid propellants to be used on Space Station. Because of the quantities involved and the frequency of extra vehicular activity (EVA), some Space Station contamination will occur. An air lock will likely serve as a decontamination station and will contain a propellant monitor. If elevated propellant concentrations are detected in the air lock, then that atmosphere will be dumped to space to prevent contamination of the cabin atmosphere. The air revitalization and trace contaminant control systems were not designed to handle high pollutant loads. Some  $N_2O_4$  contamination occurred on Apollo-Soyuz.<sup>10</sup>

Halon 1301 is the fire suppressant to be used on Space Station. Halon was detected on spacelab mission SL-1 and on Shuttle missions STS-3, and STS-4. The trace contaminant control system (TCCS) will only handle modest quantities. Halon degradation products are toxic and will poison the catalytic oxidizer. If a halon release occurs it will be necessary to vent the cabin air to space and repressurize.

Methane is a metabolic product which usually accumulates as each mission progresses. It will likely be the contaminant of greatest concentration. The Bosch  $CO_2$  reduction system, a candidate for the air revitalization system (ARS), will produce large quantities of methane. A high temperature catalytic oxidizer will be required to keep  $CH_4$  concentrations below 1 ppm.<sup>12,13</sup>

$CO$ , a product of incomplete combustion, may be released from metabolic processes, smoldering of carbon filters, or fire. The Bosch  $CO_2$  reduction system produces  $CO$  and the potential for rapid accumulation exists, if not removed by the trace contaminant control system.<sup>12,13</sup>

Ammonia ( $NH_3$ ), a product of metabolism will be released from urine processing, and it is probably a degradation product of the solid amine resin proposed for the ARS.<sup>14</sup> Phosphoric acid

impregnated charcoal filters can remove  $\text{NH}_3$ .

Hydrogen ( $\text{H}_2$ ) will be produced by electrolysis and used in  $\text{CO}_2$  reduction by both the Bosch and the Sabatier processes.<sup>12,13,15</sup> A pressure gradient will be used to minimize the likelihood of explosive mixtures from developing, if a leak occurs.

Toluene was detected on a number of missions. On Shuttle mission STS-2, toluene approached the SMAC value in one sample. Subsequent analyses indicated that for the sample, the additive toxicity hazard index for systemic poisons was exceeded by 1.22 times, with toluene the major constituent.<sup>10</sup> Toluene is also a contaminant which off gases from the solid amine resin of the ARS.<sup>14</sup>

Trimethylamine is a principal breakdown product of the solid amine resin of the ARS. The trimethylamine concentration has exceeded safe limits in tests of the ARS.<sup>14</sup> Because of the numerous trace organics off gassing from solid amine process a post sorbent bed such as phosphoric acid impregnated charcoal will be used.

Glutaraldehyde and silicon escaped containment on Spacelab mission SL-D1. Glutaraldehyde is a preservative and disinfectant with irritating properties. It may also be used in electrophoresis experiments on Space Station. Silicon compounds are catalyst poisons and will occur on Space Station.

Freons have been detected on all Shuttle missions.<sup>16</sup> The degradation products are corrosive, irritating, toxic, and catalyst poisons. Freon 12 will be on Space Station.

A computer model developed from Shuttle charcoal canister analysis for TCCS contaminant removal studies indicated that five contaminants may exceed SMAC values: propenal (acrolein), an irritant; benzene, a systemic poison and carcinogen; o-diethylphthalate, an irritant; propylfluorosilane, an irritant and catalyst poison; and 2-methylhexane, a central nervous system depressant.<sup>17</sup> Benzene has also temporarily exceeded SMAC values during preflight off gassing tests.<sup>16</sup>

Ethanal (acetaldehyde), ethanol, dichloromethane, and acetone have a high frequency of occurrence on shuttle missions and are likely to be present on Space Station.<sup>16</sup>

Oxidation products will be produced in the catalytic oxidizer. Post sorbent beds are necessary to prevent the release of oxidants and free radicals to the cabin air from the TCCS. Also, it has been hypothesized that secondary pollutants are important in cabin atmospheres. Trial simulations have indicated that spacecraft cabins may develop elevated  $\text{NO}_2$  concentrations and ozone ( $\text{O}_3$ ) concentrations exceeding SMAC values.<sup>18</sup> Oxidation products,  $\text{NO}_2$ ,  $\text{O}_3$ , and formaldehyde, were among the contaminants suspected of causing irritation on Shuttle flights, although particulates from biological sources were the undisputed cause of crew discomfort.<sup>19</sup>

Major metabolic products which must be removed by the ARS and the TCCS are  $\text{CO}$ ,  $\text{CO}_2$ ,  $\text{NH}_3$ ,  $\text{H}_2\text{S}$ ,  $\text{CH}_4$ , organic acids, and mercaptans.<sup>20</sup>

Foul odors have been observed on a number of missions. Many were attributable to metabolic products. An unusual odor and crew headaches occurred on Shuttle flight STS-6. Burnt wire insulation from an electrical short was the suspected causal agent.<sup>10</sup> Electrical fire can produce a number of noxious agents including halogenated organics, benzene derivatives, nitriles, and cyanates.<sup>21</sup> Space Station design must be able to handle such contingencies either through the TCCS or a smoke removal unit,<sup>22</sup> without having to rely on venting the cabin air to space and repressurizing.

For fire safety concerns, Halon 1301 will be used for fire suppression, followed by venting cabin air to space and repressurizing. Smoke detectors are an integral part of the fire detection and suppression system. To protect from toxic combustion products, infrared monitors are recommended for CO, hydrogen fluoride (HF), and hydrogen cyanide (HCN).<sup>23</sup>

Volatiles will be released to the atmosphere from electrolysis and from phase change urine processing. Carboxylic acids and phenols will be major contaminants.<sup>24,25</sup> Iodination products from the water disinfection process may cross the air/water interface and permeate the entire life support environment. The identity of these products, their expected concentrations, and their medical effects are largely unknown.<sup>26</sup> However, I suspect the byproduct concentrations and effects of iodination are less than those resulting from chlorination.

#### SPACE STATION CONFIGURATION

Space Station is designed to operate for extended periods, 90 plus days, without resupply. A regenerative, nearly closed loop life support system will be required to minimize resupply logistics and costs. Overboard disposal of wastes and venting of gases to space will be minimal. All waste materials will be treated and recycled. The concentrated wastes will be stabilized and stored for ground disposal. The expected useful life of the station is decades and a diversity of materials will be brought aboard for experimental or manufacturing purposes. The likelihood of cabin contamination is great. Cabin contamination can occur from a number of sources: biological waste production, material off gassing, process leakage, accidental containment breach, and accumulation due to poor removal efficiencies of the purification units.

The Space Station, "Freedom," will have four modules: the U.S. Laboratory (USLAB); the U.S. Habitation module (USHAB); the Japanese Experimental Module (JEM); and the European Space Agency (ESA) module, Columbus. The modules are connected by four resource nodes. Two airlocks and a logistics module are connected to the resource nodes. Each module will have an independent Environmental Control Life Support System (ECLSS), complete with a Trace Contaminant Control System (TCCS). The U.S. modules will have four Air Revitalization Systems (ARS), two in each module. Each ARS is designed to support four crew members; One ARS at a

time will operate in each module.

The ARS will provide ventilation to each module and node but not to the airlocks. Intramodule circulation will approximate near perfect mixing. Intermodule air exchange design is 130 cubic feet per minute (CFM) through 4-4.5 inch ducts.<sup>27</sup> The ventilation design is primarily based on: heat transfer and humidity control to maintain crew comfort; and O<sub>2</sub> supply and CO<sub>2</sub> removal requirements to maintain a healthful atmosphere.<sup>28</sup> The air exchange rate will be 1-2 years, achieved through air loss from leakage and airlock EVA.

#### ESA Columbus Module

ESA Columbus monitoring requirements are based on the fact, that the types of contaminants and their buildup characteristics are not precisely known, and that safety will require the monitoring of all contaminants permanently.<sup>29</sup> Hence, ESA is developing, for Columbus, a GC/MS to monitor N<sub>2</sub>, CO<sub>2</sub>, O<sub>2</sub>, H<sub>2</sub>O, plus a lengthy target list of trace contaminants. The contaminant list includes 15 alcohols, 5 aldehydes, 12 aromatics, 11 esters, 3 ethers, 21 halocarbons, 37 hydrocarbons, 8 ketones, and 11 miscellaneous compounds, such as NH<sub>3</sub>, CO, H<sub>2</sub>S, H<sub>2</sub>, O<sub>3</sub>, SO<sub>2</sub>, NO, and NO<sub>2</sub>.

#### Japanese Experimental Module

The 70 M<sup>3</sup> JEM and its 24 M<sup>3</sup> logistics module are designed for two crew members.<sup>30</sup> The JEM will have an independent ECLSS interfaced with the Space Station core. However, concentrated CO<sub>2</sub> will be returned to U.S. modules for processing, and wastewater will be returned to the U.S. modules for processing and H<sub>2</sub>O recovery. Its TCCS design is based on a contaminant load of 15.4 grams/day.<sup>31</sup>

The JEM will continuously monitor for total pressure, temperature, humidity, O<sub>2</sub>, and CO<sub>2</sub>. The JEM will rely on the U.S. modules for trace gas analysis via sample lines. A GC/MS is expected.

#### U.S. Modules

The U.S. modules will provide facilities for on-orbit repair, health maintenance, and a number of material processing and biological experiments intended to lead to manufacturing in space.

A maintenance work station will allow on-orbit repair of defective or damaged hardware. Processes likely to be required are drilling, sawing, welding, soldering, and epoxy gluing. A work bench/contaminant control console is envisioned that will collect the particulate and gaseous emissions generated in the repair process near their source.<sup>32</sup> The rack would be equipped with filters and the air recirculated with some venting to the TCCS. The work station would be a source of particulates, metal

fumes, and gases not encountered on prior missions.

The health maintenance facility will provide critical care for one individual for 28 days and outpatient care for the crew complement for the mission duration. The equipment and supply list for this facility will be lengthy.<sup>33</sup> It may be an additional source of trace contaminants, mainly sterilants.

The U.S. Laboratory will provide facilities for experiments and manufacturing.<sup>34</sup> The on board processes will generate biologicals, combustion and oxidation products, acid gases, metal and crystal fumes, and assorted lab wastes. Many of these materials are capable of adversely affecting the ECLSS subsystems by poisoning the catalyst or absorption beds, or they could appear in the humidity condensate, the potable water supply. Materials will have to be stored, then transported to the point of use, and the waste products handled. The lab racks will be contained with at least a two failure tolerant design. They will be equipped with some type of contaminant control equipment and vented to the TCCS. The lab racks should be equipped with monitors, specific for the process they contain to detect internal leaks. The chemical storage area should be monitored, and the cabin atmosphere must be sampled to alert the crew of any leak.

#### Trace Contaminant Control System

The technology base for the TCCS is good. Only limited system tests have been conducted but they have worked as predicted. The TCCS will consist of fixed bed charcoal filters, high efficiency particulate filters, and a high temperature (680 °C) catalytic oxidizer (palladium/aluminum) with pre and post sorbent beds of LiOH. There will be four units, two in each module. The air flow through each catalytic oxidizer is 2.5 CFM, or 5 CFM for the two U.S. modules.<sup>12</sup> This is only 4 air changes per day of what should be considered as fresh air. It may be too low. The TCCS will receive cabin air from the temperature and humidity control system. Purge gases from the ARS, waste water recovery, urine processing, waste reduction and storage systems, and lab racks are to be routed to the TCCS for contaminant removal. For comparison, the indoor air quality ventilation guideline is 15 CFM per person.<sup>35</sup> The guideline is intended to keep odors to an acceptable level to 80% of the visitors entering the space and it assumes that one third of the occupants are smoking at the rate of 2.2 cigarettes per hour.

#### TECHNOLOGY ASSESSMENT

A technology assessment study on monitoring systems was performed by Battelle Columbus Division for NASA.<sup>11</sup> They recommended: 1) a long path Fourier transform infrared (FTIR) analyzer for rapid detection of high risk contamination incidents, and 2) a GC/MS with mass selective or ion-trap technologies for detailed monitoring of extended crew exposures

to low (ppb) contaminations. Priority requirements for rating the candidate monitoring systems were: real time output, which is particularly important in closed environments and long duration missions; ability to detect and quantify a wide range and number of volatile compounds; ability to trigger a warning when the SMAC is approached; and the ability to monitor several modules and airlocks simultaneously. In the assessment, many instrument parameters were considered, each weighted equally: weight, volume, power requirements, sensitivity, dynamic range, response time, selectivity, growth capability, crew time, by-product generation, consumables, reliability, and maintenance. Instruments and systems considered included: CAMS, long path FTIR, matrix isolation FTIR, GC/MS, GC/photoionization detector, electrochemical devices, and chemiluminescent monitors. Long path FTIR and GC/MS appeared to be the most promising technologies for Space Station monitoring. However, several critical questions are yet to be answered: what is the target list of compounds to be monitored; which must be monitored on a continuous real time basis; at what concentration ranges; and in how many locations?

### CONCLUSIONS

Development of the monitoring plan for Space Station will be a continuing process. The monitoring system must be adaptable to accommodate new parameters and concentration ranges. All agents should be monitored that have a reasonable probability of occurrence at or above some action level, such one-half the SMAC. The analytical method relied upon must be able to quantify at action level concentrations. The basis for monitoring should be the contaminants: toxicity, quantities or production rates, removal efficiencies of the ECLSS system, and capacity to poison the ECLSS system.

Monitoring for all contaminants at multiple sites and at part per billion concentrations is impracticable and should not be attempted. Monitoring for a contaminant should not be done simply because there is an assigned SMAC. Equipment control monitors, and the monitoring of a surrogate parameter as a substitute for the etiological agent should not be relied upon when making health evaluations.

A GC/MS should be developed for Space Station. It should be a fixed instrument that continuously samples the well mixed cabin return air. It should have a sample line to each module and airlock for routine comparison of atmospheres from remote sections of the spacecraft. It should be capable of monitoring for major atmospheric components (e.g.  $N_2$ ,  $O_2$ ,  $CO_2$ ,  $H_2O$ ,  $H_2$ ,  $CO$ ,  $CH_4$ , and Halon 1301) and numerous trace gases (e.g. hydrocarbons, halocarbons, silanes, and phenols).

A LP/FTIR could detect modest levels of many compounds not readily amenable to GC/MS analyses (e.g. inorganics, acid and basic gases, oxidation products, and  $CO$ ). It could also serve as a trace hydrocarbon detector.

A portable hand held hydrocarbon detector should be

available for leak detection in remote areas of the spacecraft. A TLV Sniffer (has good response for CH<sub>4</sub>) or H-Nu photoionization type instrument can detect low ppm concentrations of hydrocarbons and should be adequate for this purpose.

Redundant monitors should be present for critical parameters: IR for CO; IR or electrochemical for CO<sub>2</sub>; and electrochemical for O<sub>2</sub>.

Particulates and fumes could be measured by either photometric or piezoelectric aerosol monitors. Sample collection and ground based analyses will be necessary for verifying mass loadings, particle size distributions, and chemical components often present particle bound (metals, semi-volatile and non-volatile compounds).

Each experiment and manufacturing process must be evaluated for possible sources of cabin contamination. The lab racks should be equipped with monitoring devices specific to the process being contained. There are many miniature photometric, IR, and electrochemical devices which could serve this need.

Finally, sample collection and preservation will have to be continued for ground based analyses, to confirm the accuracy and reliability of the on board monitoring system.

#### REFERENCES

1. Fundamentals of Industrial Hygiene. 3rd Edition. National Safety Council, 1988.
2. Code of Federal Regulations. 29 CFR 1910.1000. Office of the Federal Register. National Archives and Records Service. Washington D.C. 1988.
3. Threshold Limit Values and Biological Exposure Indices for 1988-1989. American Conference of Governmental Industrial Hygienists. Cincinnati, Ohio. 1988.
4. A Guide to Industrial Respiratory Protection. NIOSH Publication No. 76-189. U.S. Department of Health Education and Welfare, Washington, D.C. 1979.
5. A Guide to Safety in Confined Spaces. NIOSH Publication No. 87-113. U.S. Department of Health and Human Services. 1987.
6. Cason, R.M. and M.E. Koslin. "A Monitor for Atmospheric Composition and Contaminants in Closed Environments." Proceedings of the Intersociety Environmental Systems Conference, San Diego, Calif., July 1980.
7. Rice, J.E. and B.A. Pilon. "Atmospheric Monitoring for Submarine Applications." Proceedings of the Intersociety Environmental Systems Conference, San Diego, Calif., July 1980.

8. Strack, J.A. "Atmospheric Contaminant Monitoring and Control in an Enclosed Environment." Proceedings of the Eighteenth Intersociety Conference on Environmental Systems, San Francisco, Calif., July 1988.

9. Shuttle Flight Operations Manual. Volume 3: Environmental Control and Life Support Systems. Flight Training Branch, Training Division, Mission Operations Directorate. NASA, TD 304, August 1984.

10. Rockoff, L. A. "Internal Contamination Issues." Proceedings of the Seminar on Space Station Human Productivity, NASA, Ames Research Center, Moffett Field, Calif., March, 1985.

11. Buoni, C., R. Coutant, R. Barnes, and L. Slivon. "Space Station Atmospheric Monitoring Systems." Proceedings of the 37th International Astronautical Congress, Innsbruck, Austria, October, 1986.

12. Ray, C.D., K.Y. Ogle, R.W. Tipps, R.L. Carrasquillo, and P. Wieland. "The Space Station Air Revitalization Subsystem Design Concept." Proceedings of the 17th Intersociety Conference on Environmental Systems, Seattle, Washington, July, 1987.

13. Wagner, R.C., R. Carrasquillo, J. Edwards, and R. Holmes. "Maturity of the Bosch CO<sub>2</sub> Reduction Technology for Space Station Application." Proceedings of the 18th Intersociety Conference on Environmental Systems, San Francisco, Calif., July, 1988.

14. Wood, P.C. and T. Wydeven. "Stability of IRA-45 Solid Amine Resin as a Function of Carbon Dioxide Absorption and Steam Desorption Cycling." Proceedings of the 17th Intersociety Conference on Environmental Systems, Seattle, Washington, July, 1987.

15. Boehm, A.M., C.K. Boynton, and R.K. Mason. "Regenerative Life Support Program Equipment Testing." Proceedings of the 18th Intersociety Conference on Environmental Systems, San Francisco, Calif., July, 1988.

16. Coleman, M.E. "Summary Report of Post Flight Atmospheric Analysis for STS-1 to STS 41C." NASA-JSC Memorandum SD4-84-351, January, 1985.

17. Schwartz, M.R. and S.I. Oldmark. "Analysis and Composition of a Model Trace Gaseous Mixture for a Spacecraft." Proceedings of the 16th Intersociety Conference on Environmental Systems, San Diego, Calif., July, 1986.

18. Brewer, D.A. and J.B. Hall Jr. "A Simulation Model for the Analysis of Space Station Gas-Phase Trace Contaminants." Acta Astronautica. Vol. 15, No. 8, pp. 527-543, 1987.

19. Brewer, D. A. and J.B. Hall Jr. "Effects of Varying Environmental Parameters on Trace Contaminant Concentrations in the NASA Space Station Reference Configuration." Proceedings of the 16th Intersociety Conference on Environmental Systems, San Diego, Calif., July, 1986.
20. Poythress, C. "Internal Contamination in the Space Station." Proceedings of the Seminar on Space Station Human Productivity, NASA, Ames Research Center, Moffett Field, Calif., March, 1985.
21. Nulton, C.P. and H.S. Silvas. "Ambient Air Contamination-Characterization and Detection Techniques." Proceedings of the Seminar on Space Station Human Productivity. NASA, Ames Research Center, Moffette Field, Calif., March, 1985.
22. Birbara, P.J. and J.T. Leonard. "A Smoke Removal Unit." Proceedings of the 17th Intersociety Conference on Environmental Systems, Seattle, Washington, July, 1987.
23. Cole, M.B. "Space Station Internal Environmental and Safety Concerns." In Spacecraft Fire Safety. NTIS HC A07/MF A01. 1987.
24. Dehner, G.F. and D.F. Price. "Thermoelectric Integrated Membrane Evaporation Subsystem Testing." Proceedings of the 17th Intersociety Conference on Environmental Systems, Seattle, Washington, July, 1987.
25. Fortunato, F.A. and K.A. Burke. "Static Feed Electrolyzer Technology Advancement for Space Application." Proceedings of the 17th Intersociety Conference on Environmental Systems, Seattle, Washington, July, 1987.
26. Sauer, R.L., D.S. Janik, and Y.R. Thorstenson. "Medical Effects of Iodine Disinfection Products in Spacecraft Water." Proceedings of the 17th Intersociety Conference on Environmental Systems, Seattle, Washington, July, 1987.
27. Davis, R.G. and J.L. Reuter. "Intermodule Ventilation Studies for the Space Station." Proceedings of the 17th Intersociety Conference on Environmental Systems, Seattle, Washington, July, 1987.
28. Reuter, J.L., L.D. Turner, and W.R. Humphries. "Preliminary Design of the Space Station Environmental Control and Life Support System." Proceedings of the 18th Intersociety Conference on Environmental Systems, San Francisco, Calif., July, 1988.
29. Leiseifer, H.P., A.I. Skoog, and H. Preiss. "Columbus Life Support System and its Technology Development." Proceedings of the 16th Intersociety Conference on Environmental Systems, San Diego, Calif., July, 1986.

30. Shiraki, K., H. Fujimori, and A. Hattori. "Environmental Control and Life Support System for Japanese Experiment Module." Proceedings of the 17th Intersociety Conference on Environmental Systems, Seattle, Washington, July, 1987.

31. Yoshimura, Y., K. Manabe, N. Kamishima, M. Minemoto, S. Hatano, T. Etoh, and H. Iida. "Study of Trace Contaminant Control System for Space Station." Proceedings of the 18th Intersociety Conference on Environmental Systems, San Francisco, Calif., July, 1988.

32. Junge, J. "A Maintenance Work Station for Space Station." Proceedings of the 16th Intersociety Conference on Environmental Systems, San Diego, Calif., July, 1986.

33. Harvey, W.T., S.M. Farrell, J.A. Howard Jr., and F. Pearlman. "Space Station Health Maintenance Facility." Proceedings of the 16th Intersociety Conference on Environmental Systems, San Diego, Calif., July, 1986.

34. Perry, J.P. and W.R. Humphries. "Process Material Management in the Space Station Environment." Proceedings of the 18th Intersociety Conference on Environmental Systems, San Francisco, Calif., July, 1988.

35. Janssen, J.E. and D.T. Grimsrud. "Ventilation Standard Draft Out for Review." ASHRAE Journal. pp. 43-45. November, 1986.

N89 - 20066

MODEL FORMULATION OF NON-EQUILIBRIUM GAS  
RADIATION FOR HYPERSONIC FLIGHT VEHICLES

Final Report

NASA/ASEE Summer Faculty Fellowship Program -- 1988

Johnson Space Center

Prepared By: Ing Chang, Ph.D.  
Academic Rank: Associate Professor  
University & Department: Prairie View A&M University  
Department of Mechanical Engineering  
Prairie View, Texas 77446

NASA/JSC

Directorate: Engineering  
Division: Advanced programs Office  
Branch: Aeroscience  
JSC Colleague: C.P. Li, Ph.D.  
Date Submitted: September 27, 1988  
Contract Number: NGT 44-005-803

ORIGINAL PAGE IS  
OF POOR QUALITY

## ABSTRACT

Several radiation models for low density nonequilibrium hypersonic flow are studied in this report. It is proposed that these models should be tested by the 3-D VRFL code developed at NASA/JSC. A modified and optimized radiation model may be obtained from the testing. Then, the current VRFL code could be expanded to solve hypersonic flow problems with nonequilibrium thermal radiation.

## INTRODUCTION

When hypersonic flight vehicles, such as Aeroassisted Orbiter Transfer Vehicles (AOTVS), traveling in low density environment, strong bow shocks occur adjacent to the vehicles. Behind the shock front, the flow involved could be at thermal and chemical nonequilibrium. At the extreme high temperatures, the effect of nonequilibrium gas radiation may be prominent. Determining the magnitude of radiation in this environment is considered to be one of the most important problems associated with AOTVS (1-3).

There are many research works dealing with thermal and chemical nonequilibrium hypersonic flow, some of the recent papers are given in the references (4-10). Although the nonequilibrium radiation is part of the nonequilibrium processes, it is usually assumed as equilibrium radiation or neglected completely in the analysis. The relatively few work in nonequilibrium radiation research will be mentioned in the later sections.

The purpose of this report is to give a brief discussion of nonequilibrium radiation in low density hypersonic flow, with various models for radiative interaction, numerical solution methods, and some of the available codes. In the next section, the solution methods used for the general nonequilibrium hypersonic flow will be discussed, this is followed by the section dealing with radiation.

## THERMAL AND CHEMICAL NONEQUILIBRIUM HYPERSONIC FLOW

There are several numerical methods which can be used for the nonequilibrium hypersonic flow (11). The Direct Simulation Monte Carlo method (12,13), based on probabilistic collision model, requires excessive amount of computer time to solve a two-dimensional flow problem so it may be not able to solve a three-dimensional flow problem. The Viscous shock Layer method (14, 15), which used the simplified conservation equations in the shock

**ORIGINAL PAGE IS  
OF POOR QUALITY**

layer, fails to predict the flow for the whole vehicle. To solve the complete flow field around a hypersonic vehicle, the full Navier Stokes equations should be used.

#### Full Navier-Stokes Equations

In order to calculate radiation from the nonequilibrium region of the flow, one needs to know the thermal and chemical state of the gas in the region. This includes the nonequilibrium species concentrations and various temperatures to express the different energy modes of the species. These state variables can be calculated by simultaneously solving a set of mass, momentum, and energy conservation equations, the so-called full Navier-Stokes equations (2,16,17). The equations consist of one mass conservation equation for each species, one overall momentum equation, and several energy equations. The number of energy equations should equal to the number of temperatures which characterize the various energy modes of the species. For example, when using a two-temperature model of one electron temperature and one heavy-particles temperature, the two energy equations could be one electron energy conservation equation and one overall energy conservation equation. If the two-temperature model used involving one electron-vibrational temperature and one rotational-translational temperature, the two energy equations could be one electron-vibrational energy conservation equation and one overall energy conservation equation. For a three-temperature model of an electron temperature, a vibrational temperature for all species, and a translational-rotational temperature for all species, then there are three energy conservation equations for them. In the literature, the thermal nonequilibrium model could involve more than three temperatures (18), then the computational time has to increase greatly.

In the overall energy equation, there is a term represents the contribution of thermal radiation. This term is expressed by a divergence of a radiation heat flux vector. To evaluate this radiation heat flux vector, a radiative transport equation is needed. The coupling of these equations makes the nonequilibrium flow problem very difficult to solve. The radiative transport equation is discussed in the later section.

#### Rate Equations

To formulate the full Navier-Stokes equations for nonequilibrium flow, various rate expressions for source functions should be specified. In the mass conservation equation for each species, the source term is the mass production rate of species due to chemical reactions. The chemical reaction rate is a function of the various temperatures used for species. In the energy equation involving translational - vibrational energy

exchange, the relaxation time in the rate equation is of the Landau - Teller type (2,17). For the electron-vibrational energy exchange, the relaxation time expression is a modification of the Landau-Teller expression (2,19).

## NONEQUILIBRIUM GAS RADIATION

As mentioned in the earlier section, to include a radiation energy term in the overall energy equation, an additional equation is needed, this is the equation of radiative transfer. The coupling between this radiative transfer equation and the energy equation in gas dynamics is of a complicated integral-differential form. When nonequilibrium requirement is included, it adds more complexity to the problem. In the following sections, the basic equation, various simplification and solution methods are briefly discussed.

### The Equation of Radiative Transfer

The equation of radiative transfer is a continuity equation for the number density of photons in a specific solid angle and at a specific frequency (17). The equation can be formulated in terms of intensity, emission and absorption coefficients. In a nonequilibrium environment, a knowledge of the emission and absorption coefficients is a requisite to a careful solution of the equation of transfer. The discussion of various spectral models for emission and absorption coefficients is in the next section. Since the radiation transport is coupled with flow equations, for a formal solution procedure, one must rely on an iterative method to solve the system of flow equations with radiation transport.

### Radiation Transport Models

There are a number of spectral models for radiative interaction computations in the literature. These models are briefly outlined in the following:

Line By Line Transport Calculation Park, etc (3,20,21) has developed a model to calculate radiative properties for nonequilibrium air. The model assumes that the radiation transport is not strongly coupled with the flow equations, so the heavy-particle translational temperature, vibrational temperature, electron temperature, and species mole fractions are solved independently from flow equations. These items will be used as inputs to calculate the number densities of internal state of atoms and molecules by using the so called quasisteady-state assumption. A detailed line-by-line calculation is performed to

ORIGINAL PAGE IS  
OF POOR QUALITY

evaluate emission and absorption characteristics independently, using the calculated, nonequilibrium, excited internal state density values. This model requires a tremendous amount of computer time and memory space for the required computations. It becomes even more so when four vibrational temperatures were used for the gas (18). It is very difficult to apply this model to a general 3-D hypersonic flow problem.

The step-Function Absorption Coefficient Model A step function model has been developed to compute the radiation absorption coefficients by Zoby, Sutton, etc. (22). In general, the model includes, free-free, bound-free, and bound-bound transitions of various atoms and molecules. The inputs to the model are species number density and temperature of the flow field. The frequency dependence of the absorption coefficient is represented by a set of steps, these steps has a fixed, but not necessarily equal widths chosen to resolve the detailed spectra adequately. The total absorption coefficient of a particular step is a summation of all the average absorption coefficient of individual energy transitions in that step. The average absorption coefficient of an individual transition can be obtained by integration the absorption coefficient over frequency in that step. When the step width is properly chosen, there should be analytic expressions to approximate the absorption coefficient in that step for the purpose of integration. Thus the frequency dependence for each step becomes a set value, and the total absorption coefficient for that step can now be expressed in terms of species number density and temperature. When this is done, the radiative transfer equation is well expressed in each frequency step. In Reference 22, a 58 steps were used to cover the frequency range from 0 to 17 ev.

Although this model is mainly applied to equilibrium radiation, it includes the frequency dependent absorption coefficient (22,23,24). In addition, with some modifications, this method can be applied to nonequilibrium radiation (25).

There is a similiar but simpler approximation called the band approximation model for absorption coefficient (26,27). This model is also for equilibrium radiation and usually is a two-band approximation model.

Parametric Radiation Model The nonequilibrium radiation environment corresponds to conditions where population densities of the energy levels deviate from the equilibrium Boltzmann distributions based on a single temperature. At sufficiently low densities, the probability of a radiative transition becomes comparable with the probability of a corresponding collisional transition. Unless the gas is optically thick, the emission of a photon is not balanced by its inverse. Consequently, the

population distribution among the energy levels departs from the predicted by the Boltzmann equation (25,28). In the formulation of the radiative transfer equation, Tiwari, etc. (25,29) devised a model that the absorption coefficient of different species is calculated by using the 58-step absorption model developed by Sutton (22), while the major nonequilibrium effect will enter through the source function in the equation. For a multilevel energy transitions model, this is done by introducing a parameter  $n$  which represents the ratio of the collisional relaxation time and the radiative lifetime of the first excited state, and the higher level energy transitions can be related to this parameter. The nonequilibrium radiation become important for conditions when  $n = 0$  (1). The key to this model is the determination of collisional relaxation time between various particles and radiative lifetime of the excited states. Additional comparison and analysis are needed for this model.

Other Radiation Models There are a number of other radiation models in the literatures (30-32). Further investigations and testings are needed for these models. In the field of nonequilibrium hypersonic flow with thermal radiation, it is desirable to have more theoretical analysis, computational code developments, and experimental data collections.

#### Integration of Transfer Equation

To obtain the radiant heat flux vector in the energy equation, the radiative transfer equation needed to be integrated spectrally as well as spatially. This is very difficult to perform on a flow field over a whole vehicle. For an approximation, the tangent slab approximation technique is commonly used (20,22,23,29). This approximation implies that the radiative energy transfer along the body is negligible in comparison to that transferred in the direction normal to the body. Other approximation such as optical thin or/and optical thick assumptions may also be used to solve the problem (18). In general, additional works are needed in the area of fast spatial integration of transfer equation.

#### CONCLUSIONS

Several radiation models were found in the literature survey. These models need to be tested by some existing 3-D Navier-stokes computer codes. It is proposed that the VRFL code is used for this purpose (7). From the result of testing, comparisons can be made on accuracy of prediction, requirement of computer time and memory space, and generality of the model. To save computer time and to facilitate easy implementation, these models should be

ORIGINAL PAGE IS  
OF POOR QUALITY

simplified but retained their essential features. In the computer tests, the flow field may be limited to the stagnation region of a symmetric blunt body for tangent slab approximation, use minimum number of temperatures for thermal model, and assume radiation heat transfer equation and gas dynamic equations are uncoupled. From this testing and additional analysis, a modified and optimized radiation model may be obtained.

After the optimized radiation model is incorporated into the VRFL code, the heat transfer rate to a entire hypersonic vehicle should be calculated. The radiation heat transfer equation and gas dynamic equations are still uncoupled but the tangent slab approximation is removed. The purpose of this test is to find a numerical model for the spatial integration of the radiative transfer equation.

Finally, the expanded VRFL code can be applied to vehicles such as AOTVs. For these computations, the coupling of radiation and gas dynamics should be maintained. The purpose of calculation is not only to accurately predict the heat transfer to the vehicle, but also to find the effect of radiation to the hypersonic flow field.

New radiation model and numerical solution technique may still need to be developed. At present there are still too many unanswered questions in the field of nonequilibrium hypersonic flow with thermal radiation.

REFERENCES

1. Park, C., "Radiation Enhancement by Nonequilibrium in Earth's Atmosphere," AIAA Paper 83-0410, AIAA 21st Aerospace Sciences Meeting, Reno, Nevada, Jan. 1983.
2. Lee, Jong-Hun, "Basic Governing Equations for the Flight Regimes of Aeroassisted Orbital Transfer Vehicles," Thermal Design of Aeroassisted Orbital Transfer Vehicles, ed. by H. F. Nelson, Progress in Astronautics and Aeronautics, vol.96, 1985.
3. Park, C., "calculation of Nonequilibrium Radiation in the Flight Regimes of Aeroassist Orbital Transfer Vehicles," Thermal Design of Aeroassisted Orbital Transfer Vehicles, ed. by H.F. Nelson, Progress in Astronautics and Aeronautics, vol. 96, 1985.
4. Gupta, R.N., Simmonds, A.L., "stagnation Flowfield Analysis for an Aeroassist Flight Experiment Vehicle," AIAA paper 88-2613, AIAA Thermophysics, Plasmadynamics and Lasers Conference, San Antonio, Texas, Jun. 1988.

5. Li, C. P., Wey, T. C., "Numerical Simulation of Hypersonic Flow Over an Aeroassist Flight Experiment Vehicle," AIAA paper 88-2675, AIAA Thermophysics, Plasmadynamics and Lasers Conference, San Antonio, Texas, Jun. 1988
6. Li, C. P., "Computation of Three-Dimensional Flow about Aerobrake Configurations," AIAA paper 86-0566, AIAA 24th Aerospace Sciences Meeting, Reno, NV., Jan. 1986.
7. Li, C. P., "Chemical Nonequilibrium and Viscous Flow Computation for Conic Aerobrake Bodies," Proceedings of ISCFD, Sydney, Australia, Aug. 1987.
8. Li, C. P., "Chemistry-Split Techniques for viscous Reactive Blunt Body Flow computations," AIAA paper 87-0282, AIAA 25th Aerospace Sciences Meeting, Reno, NV, Jan. 1987.
9. Prabhu, D. K., Tannehill, J. C., Marvin, J. G., "A New PNS code for Chemical Nonequilibrium Flows," AIAA paper 87-0248, Jan. 1987.
10. Zoby, E. V., Lee, K. P., Gupta, R.N., Thompson, R. A., Simmonds, A. L., "Viscous Shock-Layer Solutions with Nonequilibrium chemistry for Hypersonic Flows Past Slender Bodies," AIAA paper 88-2709, AIAA Thermophysics, Plasmadynamics and Lasers Conference, San Antonio, Texas, Jun. 1988.
11. Tam, L. T., "A Status Report on Radiation and Thermal Nonequilibrium Modeling for AFE Flowfields," Status Report NAS9-17900, Lockheed Engineering and Management Services Company, Inc., Houston, Texas, May 1988.
12. Moss, J.N., Bird, G. A., "Direct Simulation of Transitional Flow for Hypersonic Reentry condition," Thermal Design of Aeroassisted Orbital Transfer Vehicles, Progress in Astronautics and Aeronautics, vol. 96, ed. by H. F. Nelson, 1985.
13. Moss, J.N., Bird, G. A., Dogra, V. K., "Nonequilibrium Thermal Radiation for an Aeroassist Flight Experiment Vehicle," AIAA paper 88-0081, AIAA 26th Aerospace Sciences Meeting, Reno, NV., Jan. 1988.
14. Song, D. J., Swaminathan, S., Lewis, C. H., "High Altitude Effects on Three-Dimensional Nonequilibrium Viscous Shock-Layer Flows," AIAA paper 84-0304, AIAA 22nd Aerospace Sciences Meeting, Neno, NV., Jan. 1984.

**ORIGINAL PAGE IS  
OF POOR QUALITY**

15. Cheng, H. K., Wong, E., "Fluid Dynamic Modelling and Numerical Simulation of Low-Density Hypersonic Flows," AIAA Paper 88-2731, AIAA Thermophysics, Plasmadynamics and Laser Conference, San Antonio, Texas, June 1988.
16. Appleton, J. P., Bray, K. N. C., "The Conservation Equation for a Nonequilibrium Plasma," J. of Fluid Mechanics, vol. 20, Dec. 1964.
17. Vincenti, W. G., Kruger, Jr. C. H., "Introduction to Physical Gas Dynamics," John Wiley & Sons, New York, 1965.
18. Candler, G., Park, C., "The Computation of Radiation from Nonequilibrium Hypersonic Flows," AIAA paper 88-2678, AIAA Thermophysics, Plasmadynamics and Lasers Conference, San Antonio, Texas, June 1988.
19. Katsuhisa, K., "Electron and Vibrational Energy Conservation Equations for Aeroassisted Orbital Transfer Vehicles," AIAA Journal, Jan. 1987.
20. Arnold, J. O., Cooper, D. M., Park, C., Prekash, S. G., "Line-By-Line Transport Calculations for Jupiter Entry Probes," Entry Heating and Thermal Protection, ed. by W. G. Ostad, Progress in Astronautics and Aeronautics, vol. 69, 1980.
21. Park, C., "Modeling of Radiative Heating in Base Region of Jovian Entry Probe," Entry Heating and Thermal Protection, ed. by W. G. Ostad, Progress in Astronautics and Aeronautics, vol. 69, 1980.
22. Zoby, E. V., Sutton, K., Ostad, W. B., Moss, J. N., "Approximation Inviscid Radiating Flowfield Analysis for Outer Planet Entry Probes," Progress in Astronautics and Aeronautics, vol. 64, 1979.
23. Tiwari, S. N., Szema, K. Y., "Influence of Precursor Heating on Viscous Flow Around a Jovian Entry Body," Progress in Astronautics and Aeronautics, vol. 64, 1979.
24. Sutton, K., "Air Radiation Revisited," Progress in Astronautics and Aeronautics, vol. 96, 1985.
25. Tiwari, S. N., Subramanian, S. V., "Nonequilibrium Radiative Heating of a Jovian Entry Body," Entry Heating and Thermal Protection, ed. by Ostad, Progress in Astronautics and Aeronautics, vol. 69, 1980.

26. Kumar, G. N., Vachon, R. I., "Radiation and Convection Heat Transfer Interactions in the Three Dimensional Compressible Hypersonic Tubulent Boundary on a sharp Cone at an Angle of Attack," Thermal Sciences 16, vol. 1, Hemisphere Publishing Corp., New York, 1983.
27. Kumar, G. N., Vachon, R. I., "Comparision of Three Radiative Formulations for Interactions in Three Dimensional Boundary Layers," AIAA Journal, vol 23, No. 1 Aug. 1984.
28. McWhirter, R. W. P., "Spectral Intensities," Plasma Diagnostic Techniques, ed. by R.H. Huddleston, etc., Academic Press, New York, 1965.
29. Tiwari, S. N., Subramanian, S. V., "Nonequilibrium Radiative Heating of an ablating Jovian Entry Probe," AIAA Journal, vol 21, Sept. 1983.
30. Carlson, L. A., "Approximations for Hypervelocity Nonequilibrium Radiating, Reacting, and Conducting Stagnation Regions," AIAA Paper 88-2672, AIAA Thermophysics, Plasmadynamics and Lasers Conference, San Antonio, Texas, June 1988.
31. Carlson, L. A., Bobskill, G. J., Greendyde, R. B., "Comparison of Vibration, Dissociation Coupling and Radiative Heat Transfer Models for AOTV/AFE Flowfields," AIAA Paper 88-2673, AIAA Thermophysics, Plasmadynamics and Lasers Conference, San Antonio, Texas June 1988.
32. Babikian, D. S., Edwards, D. K., "Radiation Heat Transfer in the Wake of a Hypersonic Vehicle," AIAA Paper 88-2634, AIAA Thermophysics, Plasmadynamics and Lasers Conference, San Antonio, Texas, June 1988

ORIGINAL PAGE IS  
OF POOR QUALITY

**N89 - 20067**

**A MODEL FOR PLASMA VOLUME CHANGES DURING SHORT DURATION  
SPACEFLIGHT**

**Final Report**

**NASA/ASEE Summer Faculty Fellowship Program--1988**

**Johnson Space Center**

<b>Prepared By:</b>	<b>John E. Davis</b>
<b>Academic Rank:</b>	<b>Assistant Professor</b>
<b>University &amp; Department:</b>	<b>Alma College Department of Exercise and Health Science Alma, Michigan 48801</b>
 <b>NASA/JSC</b>	
<b>Directorate:</b>	<b>Space and Life Sciences</b>
<b>Division:</b>	<b>Medical Sciences</b>
<b>Branch:</b>	<b>Space Biomedical Research Institute</b>
<b>JSC Colleague:</b>	<b>John B. Charles</b>
<b>Date Submitted:</b>	<b>August 12, 1988</b>
<b>Contract Number:</b>	<b>NGT 44-005-803</b>

## ABSTRACT

It is well established that plasma volume decreases during spaceflight and simulated weightlessness (bedrest). The decrement in plasma volume is thought to contribute to the orthostatic intolerance that has been observed in some crew members following spaceflight. To date, no studies have evaluated the effectiveness of fluid countermeasures of varying osmolality in the restoration of plasma volume and orthostatic tolerance in a controlled study. The overall objectives of this project were to: 1) provide a model that would rapidly and safely produce a fluid loss comparable to that which occurs during short duration spaceflight and 2) design a study that would determine the optimal drink solution to restore orthostatic tolerance and describe the mechanism(s) whereby orthostatic tolerance is restored. In the first part of the study, we used a diuretic (Lasix) as a model for the plasma volume changes that occur during short duration spaceflight. Four subjects (3 males and 1 female) participated in the study. Blood samples were drawn before Lasix administration (IV) and every 30 minutes for 3 hours after Lasix administration. Changes in plasma volume and plasma osmolality were determined for each blood sample. Urine flow was followed for 3 hours after Lasix injection. Plasma volume decreased by an average of  $11.5 \pm .78\%$  by hour 2 and then leveled off. Most of the loss of plasma occurred in the first 30 minutes after Lasix administration ( $-9.5 \pm .45\%$ ). Plasma volume was lowest at 2 hours and increased slightly between 2 and 3 hours ( $11.5 \pm .85\%$  to  $10.35 \pm .95\%$ ). This might be a result of fluid shifting from other fluid compartments (interstitial and intracellular) into the plasma. There was a slight decrease in plasma osmolality during the three hours following lasix ( $286.3$  to  $283.6 \pm .85$  mosm), thus the fluid loss was primarily isotonic. Urine flows peaked at approximately 40-60 minutes after lasix ( $14.7 \pm 2.4$  ml/min) and returned to pre-lasix values by 3 hours. These data suggest that Lasix can be used as a model for fluid loss during short duration spaceflight where other models (i.e., bedrest) are not appropriate because of the experimental design. In the second part of the study, we have designed an experiment to examine the effectiveness of fluid countermeasures in the restoration of plasma volume and orthostatic tolerance. For each subject, a euhydrated control orthostatic challenge (lower body negative pressure - LBNP) will be performed in order to establish a baseline response. A second condition will involve dehydration by Lasix administration (20 mg IV). Two subsequent dehydration experiments will be performed where LBNP will be administered following rehydration with a hypertonic saline solution and an isotonic saline solution. This will be done to produce a range of plasma volumes and osmolalities to compare LBNP responses. Plasma volume, leg volume, plasma osmolality, forearm blood flow, heart rate, blood pressure and cardiac dimensions (echocardiography) will be measured during rest, LBNP exposure, and recovery. During spaceflight, man is subjected to gravitational extremes (weightlessness to high G+ force). It will be important to document the effectiveness of countermeasures that could potentially minimize the negative effects of spaceflight.

## INTRODUCTION

The cardiovascular system adapts to the weightless environment of space. This is sometimes referred to as cardiovascular deconditioning. The extent of the deconditioning and the mechanism of deconditioning vary with the individual and the duration of spaceflight. The effects of the deconditioning are manifested upon return to the earth's atmosphere. Crewmembers have experienced orthostatic intolerance when they are re-exposed to the +1 G environment. Symptoms include tachycardia, a reduction in systolic blood pressure and pulse pressure, and a reduction in exercise capacity. In order to maintain crew safety, it is necessary to understand the mechanisms that underlie this cardiovascular deconditioning and find potential countermeasures to minimize its negative effects.

During spaceflight, there is a redistribution of fluid from the lower extremities to the head and upper body. It is thought that this central blood volume expansion results in a suppression of antidiuretic hormone (ADH) and a potentiated diuresis. This has been reported to decrease plasma volume by 8-15% (13). Current techniques to simulate the fluid shifts during spaceflight are limited to bedrest and water immersion. Both of these are time consuming and expensive and not appropriate for those studies that involve repetitive testing.

As a result of the plasma volume loss, the central hypervolemia that occurs initially is compensated for and thus an individual is well suited to a weightless environment. However, upon reentry crewmembers are exposed to 1.2 to 2 +Gz, and at this point due to the hypovolemia, some individuals experience cardiovascular instability. This is characterized by decreased blood pressure, decreased brain blood flow and in some cases syncope. The potential for these to occur continues and is further manifested upon egress from the spacecraft.

Countermeasures are currently being developed to minimize the cardiovascular deconditioning that occurs during spaceflight. These include exercise, lower body negative pressure (LBNP) during spaceflight, and fluid ingestion. In short duration flight it is thought that the primary factor involved in deconditioning is the loss of plasma volume. In longer duration flights, there could be venous compliance or baroreceptor sensitivity changes that contribute to the cardiovascular deconditioning.

Fluid countermeasures have been used for the past 15 years in both the Russian and US space programs. The initial countermeasures included water and saline in the form of bouillon to offset the natriuresis and diuresis that were thought to occur in response to the headward fluid shift. Since that time the makeup of the countermeasure has changed considerably including a wide variety of electrolyte supplements. The current US countermeasure is salt tablets and water to make an isotonic solution (0.9% saline) taken 2 hours before reentry. One recent study by Bungo

and Charles (2) has indicated that crewmembers who ingested a salt and water solution before reentry had fewer incidences of orthostatic intolerance upon re-exposure to gravity than those that did not use the countermeasure. This has yet to be validated in well controlled ground based research. In addition, we don't know if the mechanism for the restoration of orthostatic tolerance is the restoration of plasma volume. It is possible that there is a restoration of hydration status in other body fluid compartments and that this is restoring orthostatic tolerance. In order to evaluate the use of different drink solutions, it is first necessary to first develop a model for spaceflight fluid loss that can be used in a repeated treatment design.

#### Purpose of Research

The overall objectives of this study were: 1) develop a ground based simulation of changes in plasma volume during short duration spaceflight and 2) develop a research protocol that would determine the optimal drink solution to restore orthostatic responses and describe the mechanism(s) whereby orthostatic tolerance is restored.

#### HISTORICAL BACKGROUND

The fluid shift that occurs during spaceflight can be as much as 2 liters. This shift is thought to increase central venous pressure at least transiently, although this remains quite controversial (15). Atrial receptors are sensitive to changes in atrial pressure. Increasing stretch on these receptors decreases output to the anterior pituitary and thus decreases the output of ADH (6). This in turn decreases the reabsorption of water in the distal tubule of the kidney. In the non-human primate, this mechanism has been questioned (8). There was no suppression of ADH in response to volume expansion in the monkey. A study on Skylab failed to observe a diuresis. They report that urine flow was actually suppressed during spaceflight (16).

Recent evidence suggests that the increase in CVP is transient and might not be sufficient to stimulate a suppression of ADH (21). However, it might be sufficient to stimulate an increase in the release of ANF. There are several reports of increased levels of ANF in response to central volume expansion by head-out water immersion (4, 22). There have also been reports of a decrease in fluid intake during flight, which could also contribute to a reduction in plasma volume. Whatever the mechanism, there is a decline in plasma volume during spaceflight and simulated weightlessness. The decrease in plasma volume has been reported to be between 8-15% depending on the length of the flight (13). During simulated weightlessness (bedrest), the blood volume loss is equivalent to what is observed in spaceflight however, the plasma volume losses after horizontal bedrest are greater (13). Greenleaf et al. (10) reported an average plasma volume decrease of 15.2% following 14 days of bedrest.

Other studies have demonstrated that dehydration either as a result of heat stress (18), diuretic administration (19), or long duration exercise (11) decreases LBNP tolerance. Luft et al. (18) found that leg volume increased and arm volume decreased progressively with LBNP, but these changes were significantly less after dehydration. In a separate study, Luft et al. (19) found that LBNP tolerance was reduced following oral Lasix administration. The Lasix produced a 16.8% reduction in plasma volume. Hilton et al. (11) found that after dehydration as a result of 60 minutes of moderate exercise, LBNP intolerance occurred even after subjects were rehydrated.

To date, there have been no studies that have identified the source of the fluid loss. Total body water has been shown to decrease by 3% in short duration shuttle flights (17). It has been previously mentioned that plasma volume decreases by approximately 10-15% during spaceflight. It is unlikely that all of the fluid was coming solely from the plasma. There must be equilibration of body fluid compartments in the latter portions of the flight. It has yet to be determined how much fluid is coming from the interstitial space (ISF) and/or intracellular space (ICF).

As a result of cardiovascular deconditioning that occurs during spaceflight, there is an impairment of orthostatic responses (7,14). Several studies have indicated that orthostatic tolerance following bedrest is also impaired (10,12). In an attempt to reduce some of the potentially debilitating cardiovascular deconditioning of spaceflight and its effect on orthostatic tolerance, three potential countermeasures have been developed: exercise, LBNP during spaceflight, and fluid ingestion before reentry. The effects of exercise on cardiovascular deconditioning have yet to be documented. However, anecdotal information collected on the Skylab missions indicates that recovery from spaceflight was faster in those who participated in an inflight exercise program (20). Inflight LBNP has been used during Skylab with mixed results (14). However, a ground based bedrest study where LBNP was applied daily has demonstrated that LBNP can be used to offset some of the effects of weightlessness (12).

Several countermeasures have been used to offset the naturesis and diuresis that was observed during spaceflight. One of the first ingested countermeasures was 9-alpha-fluorohydrocortisone which minimized bed rest orthostatic intolerance (1). Hyatt and West (12) used one of the first fluid countermeasures and found that the combined use of oral saline and LBNP was quite effective in returning the heart rate and blood pressure responses to LBNP to prebedrest values, but saline ingestion alone was less effective. Greenleaf et al. (10) reported that saline consumption improves acceleration tolerance after bedrest by restoring plasma volume. Recently, Bungo et al. (2) examined the use of fluid countermeasures in several shuttle flights. They found that crew members that ingested a salt and water solution before reentry had fewer incidences of orthostatic intolerance upon re-exposure to gravity than those that did not use the countermeasure. In a recent study by Fry et al. (5), different drink solutions were ingested in order to determine the solution that produced

the greatest elevation in plasma volume and maintained the increase the longest. They found that a slightly hypertonic saline solution (1.07%) was most effective in increasing and maintaining the higher plasma volume. It was superior to a number of drink solutions including an isotonic saline solution (the present countermeasure). It has yet to be determined whether the hypertonic solution will increase plasma volume after dehydration and whether it will improve orthostatic tolerance as well as the isotonic saline fluid countermeasure.

There are various techniques to assess orthostatic responses. These include the stand test, tilt table test, human centrifugation, and lower body negative pressure (LBNP). The NASA stand test has the subject lie supine for 5 minutes and then stand quickly. Heart rate and blood pressure are then monitored for 5 minutes. This is a simple test, however, it does not rule out the confounding effects of muscle contraction and the resulting volume redistribution associated with it. By putting an individual on a tilt table and positioning the table to various inclines, the gravitational effects on the cardiovascular system can be varied. However, this test has limited usefulness in that it can only place a +1 Gz stress on the individual and it is difficult to get many physiological measurements during this type of testing. Human centrifugation is expensive and data collection is limited by the size of the centrifuge.

Lower body negative pressure allows for the gradation of the orthostatic stress while the subject is in the supine position. This technique facilitates data collection and allows for high +Gz forces to be applied to the subject. Lower body negative pressure results in venous pooling in the lower extremities. This in turn decreases cardiac filling pressure, end diastolic volume, and ultimately via the Frank-Starling mechanism stroke volume. Reflex mechanisms from both the carotid baroreceptors and cardiopulmonary receptors are activated which act to increase heart rate and total peripheral resistance to maintain arterial pressure. In addition, there is an increased sympathetic outflow and catecholamine release resulting in a reduction in end-systolic volume, thus maintaining stroke volume and cardiac output (24).

To date, no studies have evaluated the effectiveness of fluid countermeasures of varying osmolality in the restoration of plasma volume and LBNP tolerance in a controlled study. Studies of this nature have been limited by the need to do repetitive spaceflight simulations. For bedrest, there is an effect of one bedrest on the next bedrest making it impossible to look at independent effects of drink solutions in subsequent bedrests. Diuretics will cause a dehydration as a result of increasing fluid loss in the urine. Lasix is a commonly used diuretic for the treatment of hypertension that uses furosimide as its active ingredient. Furosimide acts to inhibit the reabsorption of sodium and chloride in the proximal and distal tubules, and the loop of henle. Its action is independent of any inhibitory effects of carbonic anhydrase or aldosterone (9). It is quick acting and has few side effects. For these reasons, we tried Lasix as a model for the fluid losses that occur during spaceflight.

## METHODS

### Model for Plasma Volume Changes During Spaceflight

In order to determine the total loss of plasma volume as a result of diuretic administration (Lasix), the time course of PV changes, the optimal time for fluid ingestion, and whether the loss of fluid with the diuretic is isotonic, four subjects were recruited and participated after obtaining informed consent (3 males and 1 female). Subjects entered the laboratory after a small meal in the euhydrated state. Subjects voided their bladder and were weighed. They then sat for 30 minutes to control for posture before blood sampling. An intracath was inserted in a superficial arm vein. A preliminary blood sample (5 ml) was drawn without stasis. Twenty mg of Lasix was injected IV. Urine samples were collected for 3 hours after Lasix injection for urine flow determinations. Blood samples were obtained every 30 minutes after Lasix. Hematocrit was determined by microcentrifugation, hemoglobin was determined by a Coulter Counter, osmolality was determined by freezing point depression on all samples. From the hematocrit (HCT) and hemoglobin (HG), relative changes in plasma volume were determined using the formula of Dill and Costill (3).

### Fluid Countermeasure and Orthostatic Tolerance

Eight volunteers will be recruited for this part of the study. Subjects will be given several practice LBNP tests prior to the start of the experimental phase of the study. For each subject, control LBNP test(s) will be performed in order to establish a baseline response to LBNP. In a second condition, dehydration will be produced by a 20 mg IV injection of Lasix. This will produce approximately an 11% reduction in plasma volume from euhydrated plasma volume as has been determined from the first part of this project. Two subsequent experiments will be performed where each subject is rehydrated after dehydration by Lasix with a 14 ml/kg solution of a hypertonic saline solution and an isotonic saline solution. For the dehydration and two rehydration experiments an LBNP test will be performed. This will give us a range of plasma volumes and osmolarities to compare LBNP responses. Each subject will complete all tests. Below is a summary of the 4 conditions:

1. Control LBNP
2. LBNP following dehydration
3. LBNP following dehydration and rehydration with hypertonic saline
4. LBNP following dehydration and rehydration with isotonic saline

The order of treatments will be randomized in order to eliminate any order effects on the LBNP results. All experiments will begin at the same time of day for each individual subject to eliminate any circadian effects on the findings. Testing will begin 2 hours after the ingestion of a small meal. LBNP will be performed 60 minutes after ingestion of fluids. At least 72 hours will be allowed between testing.

## LBNP Protocol

Lower body negative pressure will be applied using a plywood chamber sealed at the waist with a rubber gasket. A vacuum cleaner is then used to withdraw air out of the box. The following protocol will be used:

	Pressure (Torr)	Time at Stage(min)
Rest	0	10
Stage 1	-5	5
Stage 2	-10	5
Stage 3	-20	5
Stage 4	-30	5
Stage 5	-40	5
Stage 6	-50	5
Stage 7	-60	5
Recovery	0	10

The test will be terminated if the subject's systolic blood pressure drops suddenly (greater than 25 mmHg in 1 minute), systolic pressure reaches 80 mmHg, bradycardia occurs (drop in HR greater than 15 BPM), subject distress or subject request.

Before, during, and immediately after LBNP the parameters described below will be measured.

Leg volume is calculated from leg circumference measurements made every minute with a mercury-in-silastic strain gauge. The changes in leg volume during LBNP exposure will be used to approximate the amount of venous pooling in the legs.

A blood sample will be taken from an antecubital vein before drink ingestion, 30 minutes after drink ingestion, immediately before, and immediately after LBNP (a total of seven blood samples). Changes in plasma volume will be calculated from hematocrit (microhematocrit technique) and hemoglobin (Coulter Counter) ratios using the formula of Dill and Costill (3). A direct measurement of plasma volume will be performed once (human serum labeled 125-I) in each subject before the first experimental testing. This value will be used as the absolute plasma volume and changes will be expressed relative to that value. Plasma osmolality will be determined using freezing point depression on all blood samples. A 5 ml sample will be required to do all of the hematological analysis. Plasma levels of antidiuretic hormone (ADH), atrial natriuretic factor (ANF), aldosterone, and catecholamines (epinephrine, norepinephrine) will be measured for each blood sample. These assays require a blood sample of 15 ml.

Venous occlusion plethysmography with a mercury-in-silastic strain gauge will be used to measure forearm blood flow. The forearm vasoconstrictor response as indicated by this technique will be measured at each step of LBNP, and plotted as a function of systolic blood pressure. The slope of this response will be used to estimate vasoconstrictor responses to

baroreceptor (cardiopulmonary and sinoaortic) activation induced by the decreasing blood pressure. Heart rate will be determined using a three lead EKG. The R-R interval will be determined for each stage of LBNP. The change in R-R interval will be plotted as a function of the systolic blood pressure. The slope of this relationship will be used as an estimate of overall baroreceptor function. Blood pressure will be measured once a minute using a standard auscultatory technique. 2-D and M-Mode echocardiography (ATL 4000 S/LC ultrasound system, ATL, Botheli, WA.) will be performed at each stage of LBNP in order to assess relative changes in left ventricular dimensions with LBNP.

EMG will be used to assess abdominal and leg muscle tension in order to control for muscle tensing. The subject will be asked to maintain a resting EMG level throughout all LBNP tests.

Urine will be collected 30 minutes after drinking and immediately after LBNP in order to calculate urine flow rates.

## RESULTS

### Model for Changes in Plasma Volume During Short Duration Spaceflight

Plasma volume decreased in the first 30 minutes after Lasix administration (Figure 1). The loss of plasma volume leveled off at 2 hours ( $11 \pm .78\%$ ). By 3 hours, the plasma volume was partially restored, most likely as a result of fluid shifting from the ISF and ICF compartments into the plasma.

There were no major changes in plasma osmolality for three hours after Lasix in comparison to pre-lasix injection (Figure 2). There was a slight fall in plasma osmolality (286.3 to  $283.6 \pm .85$  mosm), indicating a slightly hyperosmotic urine.

Urine flow for each subject is displayed in Figure 3. For all of the subjects, urine flow peaked between 30 and 60 minutes, then decreased back to control level. By 3 hours, urine flow had returned to the control value.

In an attempt to determine the source of the fluid loss, the % loss of fluid from the plasma was estimated. We assumed that the plasma volume was 45 ml/kg for men and 35 ml/kg for women. These results are displayed in Table 1. Overall, for the 3 male subjects, one third of the fluid loss was from the plasma, and thus the other two thirds must have been from the ISF or ICF. The level of dehydration ( $\sim 11\%$  of the plasma) and the time (2-3 hours) would indicate that most of the fluid was coming from the ISF and not the ICF.

## DISCUSSION

Six degree head down bedrest is currently the standard simulation of weightlessness. However, it is expensive and can not be used in repetitive designs. Results from the present study indicate that Lasix produces

comparable losses of plasma volume (11%) to that observed following spaceflight (10-15%) (13). This is also comparable to oral Lasix administration. Luft et al. (19) reported plasma volume decreased 16.8% by 3 hours after oral Lasix. However, they did not follow the time course of the response. Lasix is fast acting as most of the loss of plasma volume occurred in the present study in the first 45 minutes following IV injection. By 2 hours the loss of plasma had peaked. Thus for the second part of this study, fluid ingestion should take place approximately 2 hours after lasix administration. As is the case during spaceflight the loss of fluid as a result of the lasix is primarily isotonic. Although, it is not known during spaceflight where the fluid loss is coming from, it is probably coming from the ISF first and then the ICF. In the present study, we found that approximately 1/3 of the total fluid loss was coming from the plasma and the rest from the other two fluid compartments (ICF and ISF).

There is the possibility that Lasix could influence the control of blood pressure independent of the plasma volume changes. Lasix does alter renal blood flow. It can increase or decrease renal blood flow depending on the experimental conditions (9). Furosimide has also been shown to increase atrial natriuretic factor and renin in plasma (23). However, it is difficult to separate out the hypovolemic effects of furosimide and its direct effect on the kidney and other vascular areas. This needs to be taken into account when using Lasix under these conditions. Given these limitations, Lasix can be used to simulate the plasma volume losses that occur during short duration spaceflight.

The design presented should determine the optimal drink solution to restore orthostatic tolerance and the mechanism whereby orthostatic tolerance is restored. We would anticipate, based on the study by Fry et al. (5), that the optimal drink solution should be the hypertonic saline solution if the mechanism for restoring orthostatic tolerance is the restoration of plasma volume. However, it is possible that other mechanisms might be operational. It is possible that fluid ingested moves through the plasma into the ISF and ultimately the ICF, thus restoring hydration in cells. This could in some way restore orthostatic tolerance independent of restoring plasma volume.

## CONCLUSIONS

In summary, Lasix can be used as a way of simulating the plasma volume changes that occur during short duration spaceflight. The total loss of plasma is comparable to spaceflight. Lasix is fast acting, and has relatively few side effects.

The present design for evaluating the optimal fluid countermeasures will have important implications in restoring orthostatic tolerance and function in the latter stages of spaceflight when it is essential for safe operation of the spacecraft.

## REFERENCES

1. Bohnn, B.J., K.H. Hyatt, L.G. Kamenetsky, B.E. Calder and W.M. Smith. Prevention of bedrest induced orthostatism by 9 alpha-fluorohydrocortisone. Aerospace Med. 41: 495-499, 1970.
2. Bungo, M. W., J.B. Charles, and P.C. Johnson. Cardiovascular deconditioning during space flight and the use of saline as a countermeasure to orthostatic tolerance. Aerospace Med. 56: 985-990, 1985.
3. Dill, D.B., and D.L. Costill. Calculation of percentage changes in volumes of blood, plasma, and red cells in dehydration. J. Appl. Physiol. 37:247-248, 1974.
4. Epstein, M., R.D. Lutzenheiser, E. Friedland, R.M. Acker, M.J. Camargo, and S.A. Atlas. Relationship of increased plasma atrial natriuretic factor and renal sodium handling during immersion induced central hypervolemia in normal humans. J. Clin. Invest. 79: 738-745, 1987.
5. Fry, M. Unpublished Results, 1988.
6. Gauer, O.H., and J.P. Henry. Neurohumoral control of plasma volume. Int. Rev. Physiol. 9:145 -190, 1976.
7. Gazenko, O.G., A.M. Genin, A.D. Yegorov. Summary of the medical investigations in the USSR manned missions. Acta Astro. 52: 907-917, 1981.
8. Gilmore, J.P. Neural control of extracellular volume in the human and nonhuman primate. In: Handbook of Physiology: The Cardiovascular System. Peripheral Circulation and Oxygen Blood Flow. Bethesda, MD: Am. Physiol. Soc., sect.2, vol III, part 2, chapt. 24, p. 885-915, 1984.
9. Gilman, A.G., L.S. Goodman, and A. Gilman. The Pharmacological Basis of Therapeutics. Macmillan Pub. Co., New York, 6th Edition, 1980.
10. Greenleaf, J.E., W. Van Beaumont, E.M. Bernhauer, R.F. Haines, H. Sandler, R.W. Staley, H.L. Young, and J.W. Yusken. Effects of rehydration on +Gz Tolerance after 14 days bed rest. Aerospace Med. 44:715-722, 1973.
11. Hilton, F., J.E. Davis, K.V. Levine, and S.M. Fortney. Orthostatic intolerance following acute exercise. Fed. Proc. 45: 644, 1986.
12. Hyatt, K.H., and D.A. West. Reversal of bedrest-induced orthostatic intolerance by lower body negative pressure and saline. Aerospace Med. 48: 120-124, 1977.
13. Johnson, P.C. Fluid volume changes induced by spaceflight. Acta Astro. 6:1335-1341, 1977.

14. Johnson R.L., G.W. Hoffler, A.E. Nicogossian, S.A. Bergman, and M.M. Jackson. Lower body negative pressure: third manned Skylab mission. In Johnson, R.S. and L.F. Dietlein (eds) Biomedical Results From Skylab. Washington, D.C.: NASA, p. 284-312, 1977.
15. Kirsch, K.L., L. Rocker, and F. Haenel. Venous Pressure in Space. In: Scientific Results of the German Spacelab Mission D1. Norderney, Germany, p. 500-503, 1986.
16. Leach, C.S. and P.C. Rambaut. Biochemical responses of the Skylab crewmen: an overview. In Johnson, R.S. and L.F. Dietlein (eds). Biomedical Results From Skylab. Washington, D.C.: NASA, p. 204-216, 1977.
17. Leach, C.S., L.D. Inners, and J.B. Charles. Changes in total body water during spaceflight. In Bungo, M.W., T.M. Bagian, M.A. Bowman, and B.M. Levitan (eds). Results of the Life Sciences DSOs Conducted Aboard the Space Shuttle 1981-1986. Houston, Texas : NASA, p.49-54, 1987.
18. Luft, U.C., L.G. Myhre, J.A. Loeppsky, and M.D. Venters. Specialized physiological studies in support of manned space flight. NASA Report # NAS9-14472, 1976.
19. Luft, U.C., J.A. Loeppsky, M.D. Venters, and Y Kobayashi. Specialized physiological studies in support of manned space flight. NASA report # NAS9-15483, 1987.
20. Michel, E.L., J.A. Rummel, C.F. Sawin, M.C. Buderer, J.C. Lem. Results of Skylab medical experiment M171-metabolic activity. In Johnson, R.S. and L.F. Dietlein (eds). Biomedical Results From Skylab. Washington, D.C.: NASA, p. 372-387, 1977.
21. Norsk, P., F. Bonde-Petersen, and J. Warberg. Influence of central venous pressure change on plasma vasopressin in humans. J. Appl. Physiol. 61: 1352-1357, 1986.
22. Pendergast, D.R., A.J. DeBold, M. Pazik, S.K. Hong. Effect of head-out water immersion on plasma atrial natriuretic factor in man. Proc. Soc. Exp. Biol. Med. 184:429-435, 1987.
23. Sakamoto, H., and F. Marumo. Atrial natriuretic peptide secretion in response to volume expansion and contraction in normal man. Acta End. 118:260-268, 1988.
24. Wolthius, R.A., S.A. Bergman, and A. Nicogossian. Physiological effects of locally applied reduced pressure in man. Physiol. Rev. 54: 566-595, 1974.

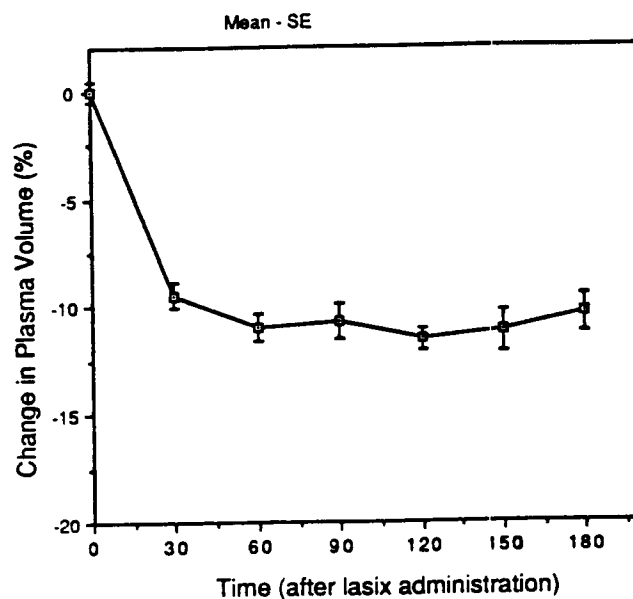


Figure 1. - Change in plasma volume as a function of time after Lasix administration. Means are plotted  $\pm 1$  SE.

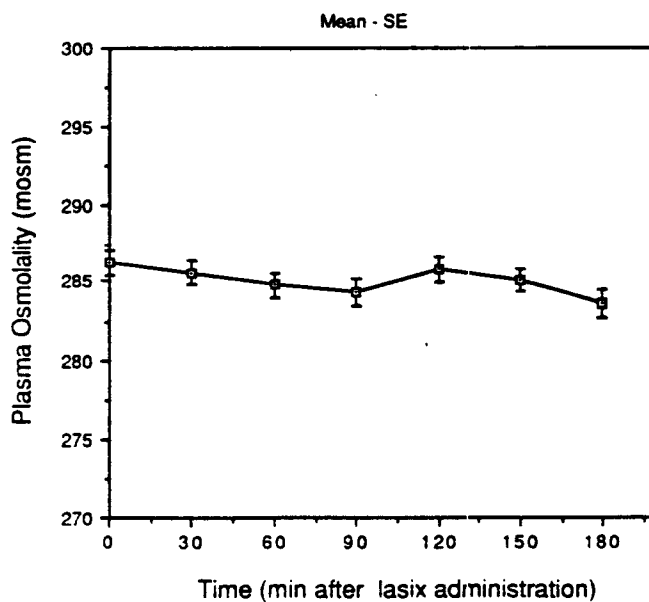


Figure 2. - Plasma osmolality as a function of time after Lasix administration. Means are plotted  $\pm 1$  SE.

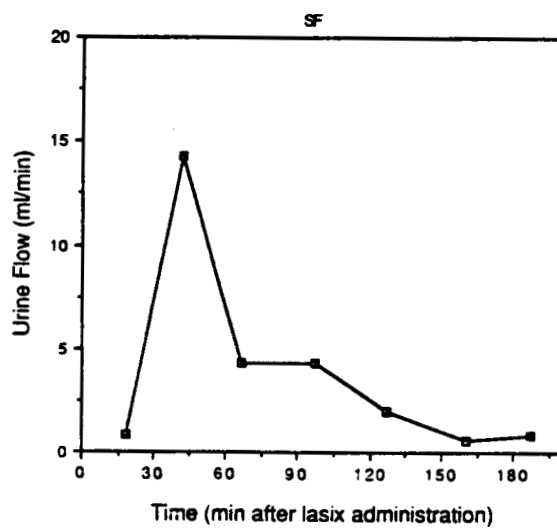
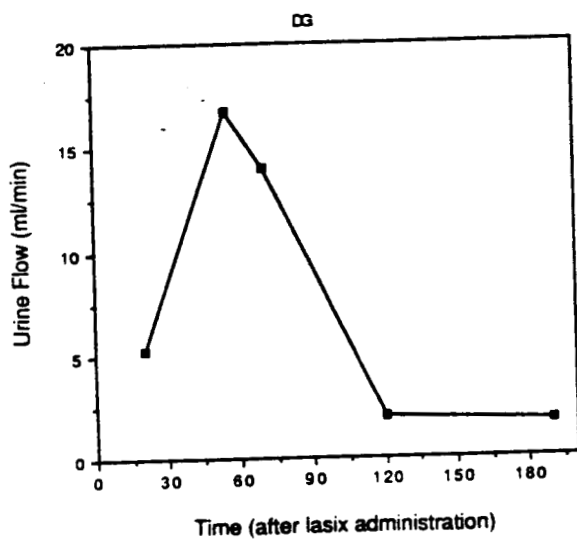
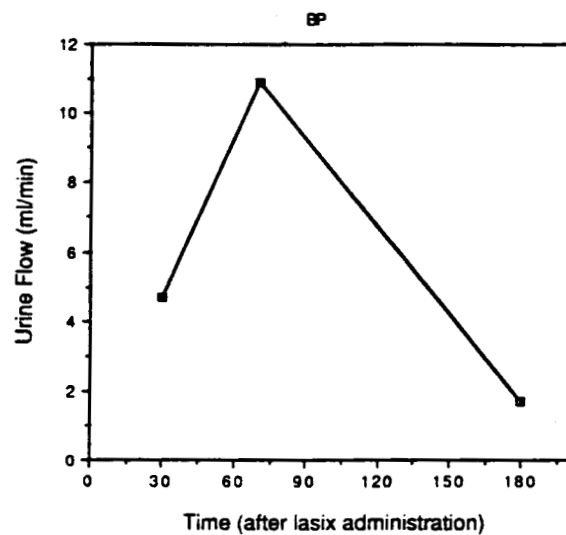
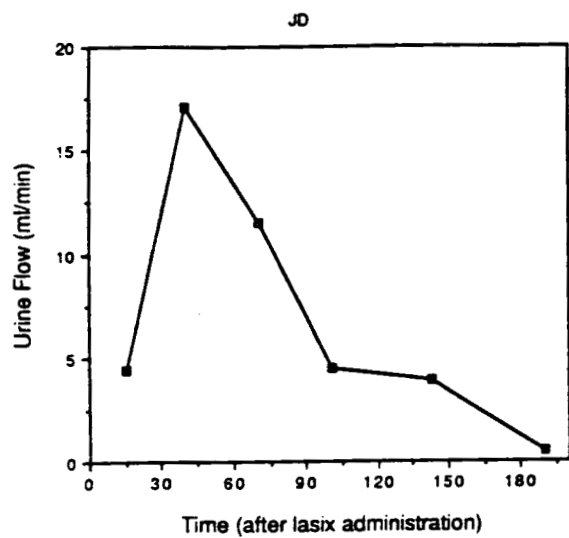


Figure 3. - Urine flow rate as a function of time after Lasix administration. Individual responses are displayed.

Table 1. - Total urine output and % loss of fluid from plasma for each subject.

SUBJECT	TOTAL URINE OUTPUT (LITERS)	% FROM PLASMA
1	1.58	34.5
2	1.13	35.8
3	1.40	31.6
4	0.72	51.7

CALCULATIONS:

$$\% \text{ LOSS FROM PLASMA} = \frac{\text{EST. PV (ML)} \times \% \text{ LOSS OF PV}}{\text{TOTAL URINE OUTPUT (L)}}$$

**AN EVIDENTIAL APPROACH TO PROBLEM SOLVING WHEN A LARGE NUMBER OF  
KNOWLEDGE SYSTEMS IS AVAILABLE.**

**FINAL REPORT**

**NASA/ASEE Summer Faculty Fellowship Program -- 1988  
Johnson Space Center**

<b>Prepared By:</b>	<b>Andre de Korvin, Ph.D.</b>
<b>Academic Rank:</b>	<b>Professor</b>
<b>University &amp; Department:</b>	<b>University of Houston - Downtown Applied Mathematical Sciences Houston, Tx. 77002</b>
<b>NASA / JSC</b>	
<b>Directorate:</b>	<b>Mission Support</b>
<b>Division:</b>	<b>Mission Planning and Analysis</b>
<b>Branch:</b>	<b>Technology Development and Applications</b>
<b>JSC Colleague:</b>	<b>Robert Lea, Ph.D.</b>
<b>Date Submitted:</b>	<b>August 12, 1988</b>
<b>Contract Number:</b>	<b>NGT 44-005-803</b>

## ABSTRACT

Some recent problems are no longer formulated in terms of imprecise facts, missing data or inadequate measuring devices. Instead, questions pertaining to knowledge and information itself arise and can be phrased independently of any particular area of knowledge. The problem considered in the present work is how to model a problem solver that is trying to find the answer to some query. The problem solver has access to a large number of knowledge systems that specialize in diverse features. In this context, feature means an indicator of what the possibilities for the answer are. The knowledge systems should not be accessed more than once, in order to have truly independent sources of information. Moreover, these systems are allowed to run in parallel. Since access might be expensive, it is necessary to construct a management policy for accessing these knowledge systems. To help in the access policy, some control knowledge systems are available. Control knowledge systems have knowledge about the performance parameters status of the knowledge systems. In order to carry out the double goal of estimating what units to access and to answer the given query, diverse pieces of evidence must be fused. We use the Dempster-Shafer Theory of Evidence to pool the knowledge bases. The present work demonstrates how pooling evidence can determine the necessary strategy to solve the problem as well as solving it. The present work deals with the path the problem solver could follow to answer the query.

## INTRODUCTION

Some recent problems are no longer formulated in terms of imprecise facts, missing data or inadequate measuring devices. Many problems today deal with knowledge and information itself and can be phrased independently of any particular area of knowledge. Problems in such diverse fields as psychology, engineering, artificial intelligence and decision making can exhibit similar formats when it comes to dealing with information and uncertainty. It is desirable to be able to handle information, not in any particular context, but in the form of general principles valid in many situations.

In [1] [2] and [8], the authors develop a model for a course of action-feedback loop that is useful when considering learning in repeated environments. This model is not applicable to the case where one-shot decisions are to be made. For such decisions, the decision maker accumulates information and after enough accumulation has taken place, certain alternatives are ruled out. This process is repeated till there is only one remaining alternative. Presumably this last alternative represents the decision to be taken. This model is studied in [9]. In this setting, it is essential to be able to update the general state of information every time an alternative is ruled out.

A powerful method in artificial intelligence is to look at certain features of a problem and try to put together the evidence coming from these features in order to recognize the corresponding pattern. In this context, feature means any indication of what the possibilities of a correct answer are. Special cases of such an approach are, for example, taken in computer vision, see [3], [13], and [14]. It is clear that with such an approach we need to have some kind of mechanism available for evidence pooling. We shall use the D-S theory [12] to carry out this pooling of evidence. In fact, the D-S theory will be used for two purposes: (a) plan the solution to a query by accessing some pertinent knowledge systems (b) define dynamically a policy that would determine what *K*Ss to access at a particular time. The problem we consider here is quite general. Given a query *Q* we look at a set of possible answers to *Q*. We of course would like to determine the precise element of the set that represents the correct answer to *Q*. In order to accomplish this, specialized knowledge systems are consulted. The execution is driven by a table that to each goal associates a sequence of control strategies. A control strategy is a set of simultaneous constraints on the desired performance characteristics of the knowledge systems to be consulted. A number of control strategies weighted by the strength of belief in each constraint is given in the table. The weight of each constraint is the degree of belief that a specific value, for some performance parameter, is optimal. This will be compounded with the uncertainty that some family of knowledge systems has the desired performance parameter value.

After the completion of each goal, new data will be read in. A control sequence defines a subcycle of the goal execution. After the completion of each subcycle, an update on the information status is performed. In order to obtain information about the knowledge systems, special knowledge systems, called control knowledge systems, are consulted. Based on the pooled evidence provided dynamically by the control knowledge systems, a policy of access is determined.

There are two common ways of representing uncertain information about propositions. The most straightforward (employed in PROSPECTOR) provides a number  $P(A|E)$  which is the probability of proposition A being true, given current evidence E. One problem with this approach is that the precision of  $P(A|E)$  is not known. In fact, Garvey et al note "A likelihood represented by a point probability value is usually an overstatement of what is actually known." Another problem with this approach is that we do not have the amount of evidence for and against A. The other way to proceed is to attach not one but two numbers with each proposition (MYCIN employs this method). The two numbers correspond to measures of belief and disbelief. This approach certainly removes one of the above criticisms. It is still true that the precision relative to the two numbers is not known. Another criticism that may be leveled is that no formal theory indicates how evidence for and against might be combined. Also, it is not clear how to detect conflict. Finally, it has been shown that results degrade quite fast as uncertainty increases. The D-S theory is used here because it avoids, for the most part, the problems outlined above.

We assume that a large number of knowledge sources is available because we do want to access independent sources and therefore do not want to access the same source more than once. In this context, "independent sources" means that the errors in these sources are independent. To approximate independence, we may force a delay before any knowledge source is reaccessed. When all the steps have been taken, a final evaluation is made to determine the answer to the query.

## CONCEPTS AND NOTATIONS

We start by defining the concept of belief structure. For a complete account of general belief structures and related material, the reader is referred to [12]. If  $X$  denotes some set and  $\tau(X)$  denotes the power set of  $X$ , a belief structure is a function  $m$  from  $\tau(X)$  into  $[0,1]$  satisfying

$$\begin{aligned} (i) \quad & m(\emptyset) = 0 \\ (ii) \quad & \sum_{A \subset X} m(A) = 1 \end{aligned}$$

Elements  $A \in \tau(X)$  for which  $m(A) > 0$  are called focal elements of  $m$ .

Now we define the Belief and Plausability measures on  $\tau(X)$

If  $B \in \tau(X)$

$$Bel(B) = \sum_{A \subset B} m(A); \quad Pls(B) = \sum_{A \cap B \neq \emptyset} m(A).$$

Suppose we have  $n$  experts whose opinions are equally respected. Assume that the opinion of the  $i^{th}$  expert is that a specified object is to be found in some set  $A_i$ . Given these  $n$  opinions we now ask the question: what is the probability that this specified object is in some set  $B$ ? Assume for ease of notation that  $A_1, A_2, \dots, A_K$  are subsets of  $B$ ,  $A_{K+1}, A_{K+2}, \dots, A_{K+\ell}$  intersect  $B$  and its complement and that  $A_{K+\ell+1}, \dots, A_n$  are subsets of the complement of  $B$ . A lower bound on the probability that the object is in  $B$  is  $K/n$ , an upper bound is  $(K+\ell)/n$ . We hence see that a possible interpretation of  $BEL(B)$  and  $Pls(B)$  is a "Lower probability and a higher probability" of  $B$ . Another way to interpret  $BEL$  and  $Pls$  is to identify the set  $B$  with the proposition: "The object is in the set  $B$ ." The truth or falsity of this proposition is determined by the available evidence, in this case, the opinion of the  $n$  experts. With this interpretation  $Bel(B)$  is interpreted to be the degree of support for proposition  $B$ .  $Pls(B)$  is the degree to which one fails to refute proposition  $B$  ( $A_1, A_2, \dots, A_{K+1}, \dots, A_{K+\ell}$  all intersect  $B$  and this constitute evidence that fails to refute  $B$ ). With this interpretation,  $Pls(\neg B)$  is interpreted as the degree to which proposition  $B$  is refuted while  $Pls(B) - Bel(B)$  is the degree of ignorance about proposition  $B$ , indeed:

$$Pls(B) - Bel(B) = \sum_{i=K+1}^{K+\ell} m(A_i)$$

and since the above  $A_i$  do intersect  $B$  and its complement, they fail to refute  $B$  and they fail to refute  $\neg B$ . Ignorance on proposition  $B$  is the degree to which one fails to refute  $B$  and  $\neg B$ . An important property of belief structures is that they allow to fuse evidence from independent sources of information. If  $m_1$  and  $m_2$  are belief structures defined on the same set  $X$  then the combination rule, see [12], states that they combine to yield a belief structure  $m_3$  on  $X$ . The focal element of  $m_3$  are given by taking all possible intersections of focal elements of  $m_1$  with focal elements of  $m_2$  and

$$m_3(C) = \sum_{A \cap B = C} m_1(A) m_2(B) / (1 - K)$$

$K$  is the measure of conflict between the two sources and comes from pairs of disjoint focal elements. In fact

$$K = \sum_{A \cap B = \emptyset} m_1(A) m_2(B)$$

If  $m_3$  is given by the formula above, we write  $m_3 = m_1 \oplus m_2$ . The formula above is true if the information sources are independent. For a very readable interpretation of the combination formula, the reader is referred to the article by L.A. Zadeh [16].

In [1], [2], and [8] the authors have defined the concept of an abstract information system. In this model, decisions are made sequentially and information is received as a consequence of the course of action taken. That is, each cycle of the decision-information acquisition loop is initiated by the decision maker (*DM*) choosing a course of action (*COA*) and the data generated by this *COA* is fed back to the *DM* which completes the cycle. This model is most relevant to situations requiring frequent decisions on a regular basis such as ordering equipment or choosing stocks. The *DM* can take the same *COA* on successive cycles or change to different *COA*, possibly returning to a *COA* chosen earlier. Uncertainty has been incorporated into this model. Different policies for selecting a *COA* on a given cycle have been discussed in [7]. The model discussed in the works cited above is not relevant to "one-shot decisions" where the *DM* uses all available information to pick a single *COA* which, once chosen, is irreversible.

In [9] the authors study situations where one-shot decisions are called for. In this context there may be many minor decisions, such as which information source to access and how to interpret information received. The *DM* constructs a set of possible *COAs* and then, based on the information received, eliminates the less promising *COAs* until only one remains. This, of course, is the *COA*

taken. The main idea in [9] is that for each fixed data vector, the *DM* creates a conditional belief structure and then after  $K$  cycles, combines these masses conditioned by the data vectors into a single mass conditioned by history. From that, the belief and plausability measures conditioned on history are formed and whenever the plausability of a  $COA_i$  falls below the belief of some other  $COA_j$ ,  $COA_i$  is eliminated from consideration. The general problem of gathering evidence through classical statistics is that, as already pointed out in the introduction, an enormous amount of data is required. In addition, with such approaches, results quickly deteriorate if uncertainty is present. Usually some version of Baye's rule of inference is used as a remedy to the large amount of micro-events present, see for example [4] and [5]. In [10] J. Pearl makes the point that a complete probability space is not required. One needs only to estimate likelihood ratios. However a large number of such likelihood ratios would still be required.

In the next section we will outline an approach that uses the DS theory to collect evidence from different cycles as in [9]. We will collect evidence by analyzing different features of the problem. Breaking a problem down to subsets of features has already been tried in diverse fields of artificial intelligence. For example for computer vision, see [3], [13], and [14]. In most of the existing approaches, information on specific features is relatively static and no update of that information is incorporated into the DS approach. In this work, an update mechanism will be built into the decision making process. Also the *DS* theory will not only be used to identify a solution of the problem but also will determine how to access the knowledge sources (*KSs*) Since access to some *KSs* might be difficult and since we will at the very least delay reaccessing a *KS*, it is particularly important to determine a viable policy for access. We will also allow different *KSs* to be multi-tasked on particular cycles or sub-cycles. Finally it is important to have a model that is potentially applicable to many situations and for this reason we have made the discussion as general as possible.

## RESULTS

The setting is as follows: Let  $Q$  denote some specific query and let  $H_Q$  be the corresponding frame of discernment, i.e.  $H_Q$  is a set of possible answers to  $Q$  and one of the elements of  $H_Q$  is the correct answer. Of course we do not know a priori which element is the correct answer. The data vector  $x$  is of the form

$$x = \left( \begin{matrix} k_1 & k_2 & \dots & k_s \\ f_1 & f_2 & \dots & f_s \end{matrix} \right)$$

where  $s$  is large and where  $f_i^{ki} \in F_i (1 \leq i \leq s)$ .  $F_i$  is a set of (not necessarily numerical) values pertaining to the  $i^{th}$  feature. For example for some  $i$ ,  $F_i$  could be different values of light intensity. These values could range in the set {low, average, fairly high, high}. For another  $i$ ,  $F_i$  could be color values such as {red, blue, green} etc... Associated with each set  $F_i$  we have a knowledge system  $KS_i$  that only reports on values in  $F_i$ . Thus for some values of  $i$ ,  $KS_i$  may report only on the value of the color. Since we would like in principle, not to access any  $KS_i$  more than once (in order to truly have independent sources of information), we assume that  $s$  is large enough to yield appropriate information to obtain an answer to  $Q$ . Given a feature value  $f_i^{ki} \in F_i$  we define a corresponding subset of  $H_Q$  by setting

$$A(f_i^{ki}) = \left\{ g \in H_Q / i^{th} \text{ feature of } g \text{ has value } f_i^{ki} \right\}$$

$A(f_i^{ki})$  can be identified with the proposition: "The  $i^{th}$  feature has value  $f_i^{ki}$ ". In fact if  $G \subset H_Q$  then  $G$  can be identified with the proposition: "The answer is in  $G$ ". Thus  $\tau(H_Q)$  can be identified with the set of relevant propositions.

As already mentioned earlier, the value of the  $i^{th}$  feature will be reported by  $KS_i$ . However, there is some uncertainty built into the reporting of  $KS_i$ . Thus for the corresponding  $i$ ,  $KS_i$  might report that the intensity is low .6, is average .3 and is fairly high .1

In general, let  $a_i^k$  be the degree of belief that the  $i^{th}$  feature value is  $f_i^k$ . Clearly

$$\sum_k a_i^k = 1, 1 \leq i \leq s.$$

We now define a belief structure  $m_i$ :

$$m_i: \tau(H_Q) \rightarrow [0,1], m_i(A(f_i^k)) = a_i^k$$

Another way to state this is to say that  $a_i^k$  is the degree of belief that any element of  $A(f_i^k)$  is possibly the correct answer. Thus to sum up: in order to get the answer to  $Q$ , we look at the values of some features. These values are reported by specialized  $KS$ s with some degree of uncertainty built into the report. It is also clear that  $m_i, 1 \leq i \leq s$  constitute independent sources of information since they deal with distinct features.

Since there is a large number of  $KS$ s to access, it is important to formulate a policy on how to access  $KS_i$ s. We base this policy on performance characteristics. Let  $P_1, P_2, \dots, P_\ell$  be sets of performance values e.g.  $P_1$  could be a set of values corresponding to the cost of access,  $P_2$  could correspond to response delays etc... To each  $P_i$  corresponds a special  $KS$ , distinct from the ones mentioned above, which knows about performance (as opposed to features). To distinguish these  $KS$ s, we call them control  $KS$ s and denote them by  $CKS$ s. Thus  $CKS_i$  has knowledge of values in  $P_i$ . This knowledge is much more dynamic than the knowledge possessed by  $KS_i$ s, since performance of the  $KS$ s strongly depends on time. We would like to have some of the  $KS_i$  run in parallel. To this end we set

$$H_p \subset \tau \{KS_1, \dots, KS_s\}$$

to be all possible combinations of  $KS$ s that may run in parallel. Some of the combinations are of course ruled out as some of the  $KS$ s might not be compatible or may have been already used. As earlier, for  $p_i^{ki} \in P_i$  define

$$A \left( p_i^{ki} \right) \in H_p$$

where each element of  $A(p_i^{ki})$  is such that its  $i$ th performance characteristic value is  $p_i^{ki}$ . For example if  $p_i^{ki}$  means the cheapest access and if  $A(p_i^{ki}) = \{KS_{12}, KS_{19}\}$  then  $KS_{12}$  or  $KS_{19}$  might be the cheapest  $KS$  to access.

In order to find an answer to  $Q$ , we will have a sequence of goals to satisfy. The process will be table-driven in the sense that each intermediate goal will correspond to a sequence of control strategies. A control strategy will be a set of performance constraints to be simultaneously satisfied. Ultimate control strategies with their respective degrees of belief may be given. For example a control strategy could be: get the cheapest access  $KS$ s (.8) and also the smallest response delays (.7). By a subcycle we mean the execution of a particular control strategy. By a cycle we mean the execution of one goal. We use the symbol  $j_k$  to refer to the  $k^{th}$  subcycle of cycle  $j$ .

On each subcycle,  $CKS_i$  assigns a mass to  $A(p_i^k)$ . Thus

$$n_i^{jk} \left( A \left( p_i^k \right) \right) = b_i^k c_i^k$$

where  $b_i^k$  marks the degree of belief that for the  $i^{th}$  performance it's the performance value  $p_i^k$  that is needed and  $c_i^k$  is the degree of belief that any element in  $A(p_i^k)$  has  $i^{th}$  performance value  $p_i^k$ .

We have  $\sum_k b_i^k = 1$  for all  $i$

$$0 \leq c_i^k \leq 1 \text{ for all } i \text{ and } k$$

and  $n_i^{j,k} (KS_1, KS_2, \dots, KS_s) = 1 - \sum_k b_i^k c_i^k$ . It is assumed that we always have

$$A(p_i^k) \neq \{KS_1, KS_2, \dots, KS_s\}$$

That is, every  $p_i^k$  determines a proper subset of  $\{KS_1, \dots, KS_s\}$ . Note that  $n_i^{j,k}$  do satisfy the properties of a belief structure that the excess mass, after  $b_i^k c_i^k$  has been assigned, is distributed over all  $KS$ s. Assume that a particular control strategy dictates that we look at particular values in performance spaces  $P_1, P_2$ , etc. ... Let  $X_1 = \{KS_{i_1}, \dots, KS_{i_p}\}$  be the set of  $KS$ s whose first performance characteristic matches  $p_1^{k_1}$ , similarly let  $X_2, X_3, \dots$  be sets of  $KS$ s whose performance characteristic matches  $p_2^{k_2}, p_3^{k_3}$ , etc... Of course this information is given by  $CKS_1, CKS_2, \dots$  at the time the driving table is consulted for the appropriate control strategy to be followed. Thus the focal elements of  $n_i^{j,k}$  are of the form  $\tau(X_1) \times H_p \times H_p \times \dots$ . The focal elements of  $n_2^{j,k}$  are of the form  $H_p \times \tau(X_2) \times H_p \times \dots$  etc...

We now use the  $D$ - $S$  theory to formulate a policy of selecting a set of  $KS$ s. Of course each  $X_i$  reported by  $CKS_i$  is assumed to be compatible with  $H_p$  i.e., if  $X_i = \{KS_2, KS_5, KS_{10}\}$  and  $\{KS_2, KS_5\} \in H_p$  but  $X_i \notin H_p$  then  $KS_{10}$  will not be considered, but running  $KS_2$  and  $KS_5$  in parallel will be considered. For example

$$n_i^{j,k} \left( \left\{ KS_2, KS_5 \right\} \times H_p \times H_p \times \dots \right)$$

will denote the degree of belief that for some value  $p_i^{j_{x_1}}$  for access cost (assuming access cost is  $P_1$ ) that value  $p_i^{j_{x_1}}$  is the best to consider compounded with the degree of belief that  $KS_2$  and  $KS_5$  have an access cost value of  $p_i^{j_{x_1}}$ . Now on subcycle  $j_k$  if we consider, for example, three performance characteristics which, for ease of notation, we assume to be  $P_1, P_2, P_3$  then we will consider the performance space

$$H_p^{j,k} = \epsilon(X_1) \times \epsilon(X_2) \times \epsilon(X_3)$$

where  $\in (X_i) (1 \leq i \leq 3)$  are the possible values of  $X_i$  (there may be several possible values of  $X_i$  due to uncertainty wherever  $p_i^{j_k}$  is the best value to consider in  $P_i$ ). If  $Y_1, Y_2, Y_3$  are possible values of  $X_1, X_2$ , and  $X_3$  then  $n_1^{j_k}$ , for example, will be induced on the set  $\{Y_1\} \times \in (X_2) \times \in (X_3)$  by

$$n_1^{j_k} \left( \left\{ Y_1 \right\} \times \in \left\{ X_2 \right\} \times \in \left\{ X_3 \right\} \right) = n_1^{j_k} \left( \left\{ Y_1 \right\} \times H_p \times H_p \times \dots \right)$$

thus, in contrast to  $H_Q, H_p^{j_k}$  (which is the set of answers of what KSs to use on subcycle  $j_k$ ) is highly dynamic and is generated by the appropriate control sequence. In general  $H_p^{j_k}$  is a set of elements of the form  $(A_1, A_2, \dots, A_{t_k})$  where  $t_k$  is the number of performance characteristics considered on a particular control strategy on cycle  $j$  and each  $A$  is an appropriate set of KSs. We now combine the masses on  $H_p^{j_k}$  by setting

$$n^{j_k} = \bigoplus_{i=1}^{t_k} n_i^{j_k}$$

The problem is to select the "best combination"  $V = (B_1, B_2, \dots, B_{t_k})$ . Note again that independence of information sources is justified as different CKSs are pooled. To get the "best combination"  $U$  we maximize

$$Bel(V) - Bel(\neg V)$$

over  $V \in H_p^{j_k}$ , and relative to the pooled belief structure  $n^{j_k}$ . The rationale for the optimization described above is to maximize the support of  $V$  over the support of its competitors. If several elements are tied for this optimization, we select one of the elements at random. If  $V$  is an optimal element we make an arbitrary selection  $b_i \in B_i, (1 \leq i \leq t_k)$ . Each  $b_i$  denotes some  $KS \in B_i$ . Now we follow the policy of running the  $b_i$  in parallel. We of course keep track of which  $b_i$  have been run and update  $H_p$  appropriately after each subcycle, so no KS will be accessed twice.

Now at the beginning of cycle  $j$  assume that the belief structure relative to subsets of  $H_Q$  is given by  $m^j$ . Of course  $m^j$  itself was obtained by pooling together the belief structures  $m_i$  where  $i$  ranges over all features considered prior to cycle  $j$ . On the first subcycle of cycle  $j$ , assume that the policy described above indicates that features  $i_1, i_2, \dots, i_t$  should be looked at. Pooling the corresponding KSs we obtain the belief structures  $m_{i_v}^{j_1} (1 \leq v \leq t)$  and we form

$$\Delta m^{j_1} = \bigoplus_{v=1}^t m_{i_v}^{j_1}$$

At the end of cycle  $j$ ,  $m^j$  is updated by

$$m^{j+1} \leftarrow m^j \oplus \Delta m^{j_1} \oplus \Delta m^{j_2} \dots \oplus \Delta m^{j_v}$$

where  $v$  denotes the number of control sequences in cycle  $j$ .

At the completion of each subcycle we check if

$$Pls(A(f_i^k)) < Bel(A(f_i^j)) \text{ for } j \neq k$$

We use the current value of  $m_j$  to do this. Of course that current value is given by

$$m^j \leftarrow m^{j-1} \oplus \Delta m^{j_1} \oplus \dots \oplus \Delta m^{j_k}$$

if we have just completed subcycle  $j_k$ .

If the inequality above is satisfied we rule out the proposition "the value of the  $i^{th}$  feature is  $f_i^k$ ". The rationale for doing this is if, for example, the upper probability for value green falls below the lower probability of some color value, say blue, then we should rule out that the color is green. Of course we could embellish this rule and require for instance that  $Bel(A(f_i^j))$ , at least initially, exceeds  $Pls(A(f_i^k))$  by some margin which could be relaxed as data accumulates. We then should perform a post-elimination update. There are many ways to do this. A straightforward method would be to distribute the mass of  $A(f_i^k)$  uniformly over the remaining  $A(f_i^j)$ . Of course if we have some additional information, the distribution may be chosen not to be uniform. We also could use the plausability of the remaining values of  $F_i$  to guide our redistribution of mass.

At the end of the cycle we have reduced the possible number of values in some of the sets  $F_i$ . We have at this point followed the actions dictated by the current goal. Now we take a new reading of the data and proceed to work in a similar manner on the next goal. A special note should be made when the control strategy is to reduce ignorance, which was defined in the previous section. Say the largest ignorance is for  $\{q_1, q_5, q_7\} \subset H_Q$ . The CKS in charge of ignorance (which is treated here as a particular feature) has access to all  $F_i (1 \leq i \leq s)$  and the corresponding  $A(f_i^{k_i})$ . To reduce the ignorance on  $\{q_1, q_5, q_7\}$  one must get hold of those KSs that do not deal with  $q_1, q_5, q_7$  simultaneously (otherwise the ignorance may continue to propagate). Thus if KSs has information about a set containing say  $q_1$  but not  $q_5$  or  $q_7$ , then KSs is a viable candidate to reduce ignorance.

Finally when all the actions corresponding to all of the listed goals have been executed we end up with some of the sets  $F_i$  reduced (e.g. value, green and red may have been eliminated from  $F_3$ , if the third feature is color value) Now to get the answer to query  $Q$  we maximize

$$Bel(q) - Bel(\neg q) , q \in H_Q$$

relative to the final value of  $m^j$ .

The rationale for the optimization defined above is of course the same as earlier. We want to maximize the support for a particular answer over its competitors.

To sum up we have used the *D-S* theory to pool the evidence given by the *CKSs* to select a policy by which we would use the appropriate *KSs* to find information relative to query  $Q$ . Then we pooled the evidence collected from the *KSs* to rule out certain propositions about feature values. Each time we ruled out a proposition we had to redistribute its mass over remaining propositions pertaining to a specific feature. When all the steps are completed we pick the answer that maximizes support over its competitors. Of course, there are other ways to pool beliefs from different sources and for a sample of these methods we refer the reactor to [6], [11], [15], and [17].

Finally, if it is desired to access *KSs* more than once, some delay must be built into reaccessing any *KS* in order to approximate the independence of information sources. Recently, theories more general than the *D-S* approach have been studied. More flexible rules than the *D-S* rule of combination have been developed. These settings will be investigated in future work.

## REFERENCES

1. Alo, R., Kleyle, R.M. and de Korvin, A. (1985), "Decision Making Mechanisms and Stopping Times in a Generalized Information System", Mathematical Modelling, Vol. 6, pp. 259-271.
2. Alo, R., de Korvin, A. and Kleyle, R. (1987), "Emergence of a Dominant Course of Action in a Generalized Feedback Loop when Goal Uncertainty is Present", Journal of the American Society for Information Science, Vol. 38, pp. 111-117.
3. Davis, L.S., Rosenfeld, A. "Computer Vision Systems". Hanson, A.R., and Riseman, E.M. eds., 101, Academic Press, New York (1978).
4. Duda, R.O., Reboh, R., et al (1979) "A Computer Based Consultant For Mineral Exploration" SRI Final Report For Project II 6415, Artificial Intelligence Ctr., SRI International, Menlo Park, California.
5. Duda, R.O., et al (1976) "Subjective Bayesian Methods For Rule-Based Inference Systems" Proc. Nat. Comp. Conf. pp. 1075-1082. AFIPS Press, Reston, Va.
6. Garvey, T.D., (1976) "Perceptual Strategies For Purposive Vision" SRI International Tech. Note No. 117, Artificial Intelligence, Ctr, Menlo Park, California.
7. Kleyle, R.M. and de Korvin, A. (1984), "Switching Mechanisms in a Generalized Information System", Cybernetics and Systems, Vol. 15, pp. 145-167.
8. Kleyle, R.M. and de Korvin, A. (1986), "A Two Phase Approach to Making Decisions Involving Goal Uncertainty", Journal of Information Science, Vol. 12, pp. 161-171.
9. Kleyle, R.M., de Korvin, A. "A Belief Function Approach To Information Utilization In Decision Making" Submitted.
10. Pearl, J. "How To Do With Probabilities What People Say You Can't" The Second Conf. on Artificial Intelligence Appl. pp. 6-11, IEEE Comp. Soc. Press Miami, Florida (1985).
11. Selfridge, P.G (1982) "Reasoning About Success and Failure In Aerial Image Understanding" Ph.D. dissertation, Computer Science Dept. Univ. of Rochester.

12. Schafer, G. (1976), A Mathematical Theory of Evidence. Princeton, NJ, Princeton University Press.
13. Tanimoto, S.L., Computer Vision Systems. Hanson, A.R., and Riseman, E.M., eds Academic Press, New York (1978).
14. Waltz, D.L. Computer Vision Systems. Hanson, A.R., and Riseman, E.M., eds Academic Press, New York (1978).
15. Yakimovsky, Y., Feldman, J.A., (1973) in Proc. Third Int. Joint Conf. on Artificial Intelligence, 580. The International Joint Conf. on Artificial Intelligence.
16. Zadeh, L.A. (1986), "A Simple View of the Dempster-Shafer Theory of Evidence and its Implication for the Rule of Combination", The AI Magazine, pp. 85-90.
17. Zucker, S.W. in Pattern Recognition and Artificial Intelligence, Chen, C.H. ed Academic Press, New York.

ORIGINAL PAGE IS  
OF POOR QUALITY

N89 - 20069

EFFECTS ON MOTOR UNIT POTENTIATION AND GROUND  
REACTION FORCE FROM TREADMILL EXERCISE

Final Report

NASA/ASEE Summer Faculty Fellowship Program--1988

Johnson Space Center

Prepared By:	Reid P. Elam, Ph.D.
Academic Rank:	Assistant Professor
University & Department:	Central Washington University Department of Physical Education Ellensburg, WA 98926

NASA/JSC

Directorate:	Space and Life Sciences
Division:	Medical Sciences
Branch:	Space Biomedical Research Institute
JSC Colleague:	Victor Schneider, M.D.
Date Submitted:	August 12, 1988

ABSTRACT

Exercise countermeasures in long duration space flight have primarily been focused on two modalities, the treadmill and bicycle ergometer. Their use in space flight has been shown to offset the deleterious effects of disuse resulting from habituation in a weightless environment. Thus, muscle atrophy, osteoporosis and neuromuscular inefficiency -- physiological consequences from prolonged exposure to microgravity -- can be controlled by use of these modalities and their proper exercise regimen. Because muscular atrophy has been most significant in the posterior calf muscles and muscles of the thigh, the use of exercise modalities that concentrate on the musculature of the lower body is warranted. The treadmill and bicycle ergometer both require similar musculature in the activities of running and cycling although the sequence of their actions and intensity of their recruitment patterns may vary. Also, the reaction between the surface of a treadmill and the lower segments in ground reaction force (GRF) measures differ from pedal to foot reactions in cycling, running GRF's being far greater than those measured in cycling. Motor unit potentiation (MUP) as measured by electromyography (EMG) is indicative of muscular actions and may be greater where stress from increased GRF is imposed. Therefore, greater MUP from treadmill exercise may be expected due to the greater GRF generated in running over cycling. This study was conducted to analyze the characteristics of MUP and GRF in treadmill exercise at the inclines of 0, 5.5 and 11 % with conjunctive speeds of 7.5, 6, and 5 MPH respectively. These speeds and corresponding inclines were set to provide equivalent physiological workloads at 12.5 METS. EMG recordings were taken from the rectus femoris and gastrocnemius of the right leg from 5 subjects. Simultaneous GRF recordings were obtained from a Delmar Avionic treadmill rigged with load cells. Measures for MUP and GRF were taken over a period containing 10 strides at steady pace. Results of ANOVA revealed no significant differences between MUP of the rectus femoris at different speed/incline settings. The same was true for the gastrocnemius. Significant differences were found between MUP of the rectus femoris and gastrocnemius at all treadmill settings ( $p < .01$ ), with the MUP of the gastrocnemius being greater. No significant relationships between MUP of the thigh and GRF, and MUP of the calf and GRF were found ( $r = .097$  and  $r = -.31$  respectively). Significant differences were found among GRF with 11% grade higher than 0% ( $p < .05$ ) and 5.5% over 0% ( $p < .01$ ). It was concluded that the gastrocnemius was more evident in EMG activity in all speed/incline settings over the rectus femoris, and that inclines from 5.5-11% produced a greater GRF's over 0%. Recommendations for future studies was made.

## INTRODUCTION

Long duration space flight has a profound effect on the human musculoskeletal and neuromuscular systems. Reported consequences from prolonged habitation in a microgravitational environment include muscular atrophy (2,4,5,7,12), bone loss (1,11,17,18) and decrements in neuromuscular performance (2,9,10,12,14,20).

These physiologic problems are evident particularly in the lower body segments and arise from disuse (1,14,15). The absence of weight-bearing stresses to musculature and their supporting structures, as experienced in 1 g, lead to these maladies (1,11,15,16). Because the use of lower body segments in locomotion is no longer a requisite for ambulation, these structures are virtually devoid of volitional input. This creates in essence, a neuromuscular shunt, and a situation similar to what is observed in denervated muscle occurs (3,6).

Exercise countermeasures have been reported to have some degree of efficacy in offsetting the debilitating effects of weightlessness (4,11,12,14,20). In particular, the exercise modalities of primary use in manned space flights have been the treadmill and bicycle ergometer since these modalities concentrate their effort-requirements to the lower body (11,15,16).

In previous studies, space flight has been shown to have an atrophying effect on muscles of the posterior calf (4,19,20) as well as the thighs and buttocks (19). Muscle strength deficits from post-flight studies have been observed in dynamometric measures of isokinetic force production of the lower limbs (14,20). Electromyographic (EMG) studies showed that muscular efficiency and fatigue onset was increased after 1-14 days in space, and shifts to higher frequencies in power spectrum analysis following flights of longer than 2 weeks duration, implied muscle deterioration (12). Studies like these have supported the postulates concerning disuse atrophy, osteoporosis and neuromuscular degradation.

Measurements comparing ground reaction or vertical forces of cycling and running have shown that ground reaction forces are greater in running than in cycling (1,15,16). Reasons for this difference center on factors of gravitational force, or forces imposed on the body and its supporting appendages. Basically, locomotion by running involves a great amount of decelerative or negative forces which provide for the storage of elastic energy to be released in subsequent positive muscular contractions (13). This was termed the stretch-shortening cycle by Komi (8) and hints at the existence of eccentric loading in phases of human locomotion (eg. running) which may be absent in non-weight bearing activities (13). Therefore, the greater reaction forces between body segments and surface impact from

the effects of gravity, the greater the requirements for muscular activity. Subsequently, it is the increased ground reaction forces that provides the implication of greater muscular activity involved in running over cycling.

The purpose of this study was to:

1. make a basic 1 g analysis of motor unit potentiation and ground reaction force during treadmill running at different speeds and inclines.
2. study any relationships between motor unit potentiation and ground reaction force during treadmill running at different speeds and inclines.
3. establish a means for assessment of individual force/muscle activation characteristics for implementation of protocol to other studies, to include those for training/testing in a weightless environment.

## REVIEW

### Motor unit potentiation.

The temperospatial aspects governing human movement are derived from the functioning of the basic components of the integrated neural and muscular systems. Together these systems comprise the operative element, the motor unit. It is the motor unit, when activated via an electrochemical sequence that controls muscular contractions and subsequent limb movement. The force and speed at which limb movements occur are dictated by several factors as follows:

1. the number of motor units recruited.
2. the frequency of firing of motor units.
3. the synchrony of motor unit firing.
4. the pattern of motor unit recruitment.
5. the type of motor units activated.
6. the inhibition of antagonist motor units.

These factors are applicable as determinants of the amount muscular activity produced during running and cycling.

EMG is a means by which the electromechanical dynamics of the neuromuscular system can be measured. Activation of muscular contractions are measured by EMG as electrical current is generated from the difference in potential sources of a dipole. The potential differences at each pole reflect the electrical activity to ensue, and the term motor unit potentiation is used to characterize this.

### Ground reaction force.

Forces that the body exerts against the earth's surface in movement, in line with Newton's third law, can be measured by low frequency output from piezoelectric sensors built into load cells. Changes in the forces exerted against the

## ORIGINAL PAGE IS OF POOR QUALITY

running surface of the treadmill are transduced into electric signals which can be calibrated to units of G force.

Often, ground reaction forces are similarly termed vertical or transient forces as the line of gravity is vertical and transmitted along the longitudinal axes of bone structures (transient) in upright ambulation.

### METHODS

Subjects were selected from among volunteers who were screened for CHD, were less than 40 years of age, had a completed Air Force Class III physical examination, and were familiar with treadmill exercise (N = 5). Following a complete orientation, subjects signed an informed consent document.

Preparation of the subjects for EMG recording, entailed the use of alcohol prep pads to cleanse and abrade the skin for placement of surface electrodes. The skin was rubbed until it became erythematous and has an electrical resistance less than 10,000 ohms.

The rectus femoris and gastrocnemius of the right leg were prepared for electrode placement. Prior to placement of the electrodes, a measurement was taken from the point of origin of each muscle to its insertion with the midpoint cited as the point of innervation. Two Ag-AgCl electrodes were placed on either side of this point so that their centers were 4 cm apart. Leads from these electrodes were inserted into a signal conditioner worn at the waist of the subject. The input EMG signals were amplified and transmitted by infrared telemetry to a signal receiver capable of receiving simultaneous signals. These signals were transferred via BNC connection to a Compaq 40 MB hard disk computer for A/D conversion and storage for later analysis.

Ground reaction forces were recorded from four load cell transducers located under the belt of a Delmar Avionics motor-driven treadmill. Analog signals were sent via hardwire BNC connection to the Compaq computer for A/D conversion and storage for future analysis.

Two signal receivers were placed on the wall on the side of the treadmill such that the field of signal acceptance would be perpendicular to the emitted signal.

Each subject was required to perform treadmill running in one session at three different speeds and inclines. Following a warm-up period including stretching exercises and 4 minutes of low intensity treadmill running, the treadmill speed was randomly (as respects order) increased to settings of 5, 6 and 7.5 MPH. These speeds were set conjunctively with grades of 11, 5.5 and 0 % respectively. These speeds and grades (incline) were selected to represent a physiological equivalency in workload at 12.5 METS (1 MET =  $\dot{V}O_2$  of 3.5 mlO<sub>2</sub>/kg/min at rest).

ORIGINAL PAGE IS  
OF POOR QUALITY

The subjects were instructed to maintain pace at 1-2 minute intervals at the prescribed speeds and inclines such that EMG and ground reaction forces could be taken for a segment of 10 strides.

After completion of the exercise, a cool-down period of 5 minutes was administered and consisted of low intensity walking at a 0% grade. This was followed by a 10 minute recovery period with the subject in a seated position.

Analysis.

Analysis of the raw EMG was performed by integration, and averaged over the ensemble recordings for the 10 strides to produce a mean value for motor unit potentiation.

Measures of ground reaction force was analyzed for the landing-support-takeoff phase of the running stride. An ensemble average of the G force was calculated over 10 strides.

Averaged values for motor unit potentiation and ground reaction forces were compared intersubjectly to note differences between variables in prescribed speeds and inclines of the treadmill. Correlations for motor unit potentiation and ground reaction force were made to analyze relationships between the two variables.

RESULTS

The results of the study are contained in Table 1, and a sample recording of G forces and EMG shown in full wave rectified form are shown in Figure 1.

Results from ANOVA showed that there were no significant differences in motor unit potentiation of the quadriceps at all speeds and incline settings ( $p < .05$ ). The same was found to be the case in differences between motor unit potentiation of the gastrocnemius at the various speed and incline settings. There was a significant difference between motor unit potentiation of the quadriceps and gastrocnemius at all speeds and incline settings ( $p < .01$ ), the mean integrated EMG (IEMG) being greater than quadricep IEMG. Significant differences existed in ground reaction forces [ $F(2,12)=12.52; p < .01$ ]. In post-priori comparison t tests, ground reaction forces generated at 5.5% grade and 6 MPH speed, and 11% grade and 5 MPH speed, were greater than at 0% grade and 7.5 MPH speed ( $p < .01$  and  $p < .05$  respectively). Ground reaction forces between 11%/5 MPH and 5.5%/6 MPH showed forces to be slightly higher at 5.5%/6 MPH, however the difference was not significant ( $p < .05$ ).

The correlation between motor unit potentiation of the quadriceps and ground reaction forces was not significant ( $r = .097$ ), nor was the correlation between gastrocnemius motor unit potentiation and ground reaction force ( $r = -.31$ ). This

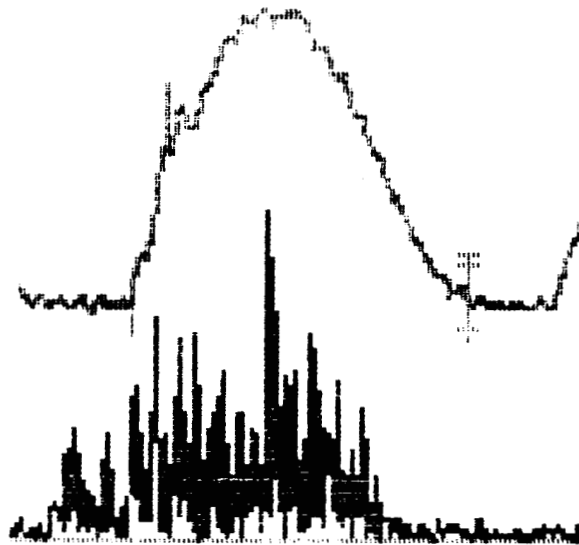
TABLE 1. RESULTS OF TREADMILL RUNNING ON MOTOR UNIT POTENTIATION (MUP) AND GROUND REACTION FORCE (GRF) FOR DIFFERENT TREADMILL (TM) SPEEDS AND GRADES. VALUES ARE MEANS (N = 5).

TM (MPH/%)		MUP (uVs)		GRF (G)
Speed/ incline	Rectus femoris	Gastroc- nemius	TOTAL	Touchdown/ takeoff
5 / 11	166.164	372.438	538.602	2.016
6 / 5.5	203.838	302.598	506.436	2.104
7.5 / 0	226.968	344.258	571.226	1.816

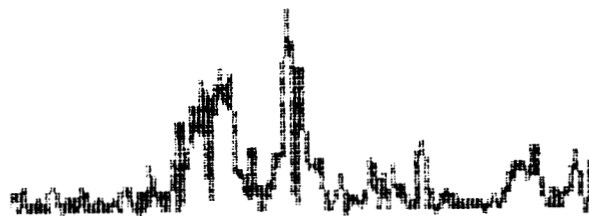
ORIGINAL PAGE IS  
OF POOR QUALITY

ORIGINAL PAGE IS  
OF POOR QUALITY

a.



b.



c.

Figure 1. Example from one subject for (a) force recording and full wave rectified EMG recording for (b) the gastrocnemius and (c) rectus femoris at 11 % grade and 5 MPH treadmill setting.

ORIGINAL PAGE IS  
OF POOR QUALITY

was for all speed and incline settings combined.

CONCLUSION

The fact that impact or collision forces exist between the surface of the treadmill and the feet of the exerciser, a vertical force is transmitted along the longitudinal axes of the bone structures. As long as this force is present, the resulting transarticular reaction forces, particularly at the talofibular/tibial and femorotibial surfaces, will require added supporting contractions of the surrounding musculature for joint stabilization. The demands imposed on the musculature of the thigh in knee stability and of the shank in ankle stability may be increased as the reaction forces become greater due to changes in speed and incline. This was not found to be the case in this study, as no significant differences were found between EMG activities of the rectus femoris at all speeds and inclines. The same was true for the gastrocnemius. Also, no correlations were found between ground reaction forces and motor unit potentiation of either the rectus femoris or gastrocnemius. Significant differences were found between potentiation of the gastrocnemius and rectus femoris at all speeds and inclines which showed that the calves are more active relative to the quadriceps in all cases analyzed in this study.

Because running requires the individual to be airborne in between strides, the resulting G forces upon impact will be greater than in walking. That increases in speed require a greater pushoff in the takeoff phase, resulting in greater airborne time, the ensuing impact may be thought to be greater, however the contrary was found to be true from the results of this study. Greater G forces were found at inclines of 11% and 5.5% over horizontal running at speeds of 5 and 6 MPH respectively versus the greater speed of 7.5 MPH at 0% grade. This was attributed to the greater horizontal impedance necessary to be exerted against the belt for pushoff in the horizontal direction. This detracts from the vertical pushoff force that was measured by the load cell instrumentation.

Greater EMG activity of the gastrocnemius at 11% grade may be attributed to the need for more forefoot running, therefore eliciting a greater stretch on the calf muscles and resulting in a stronger muscular impulse. This supports Komi's (8) assertion of eccentric factors contributing to the stretch-shortening reflex (stronger ensuing positive contraction) in running. It was interesting to note that the potentiation of the gastrocnemius at 5.5% was lower than both the 11% and 0% grades. This would show that either increased speed or inclines at 11% or higher may be capable of eliciting higher EMG responses than at lesser speeds at mid-range grades -- perhaps 1-10%.

Ground reaction forces were greater at inclines of 11%

and 5.5% over 0%, however the higher values at 5.5% over 11% indicated that there may be an optimal grade for promoting higher G forces in running at moderate speeds.

In summary, the results of this study revealed that speed and incline variation in treadmill exercise at 1 g can effect ground reaction force such that higher G forces can be produced at grades ranging from 5.5 to 11%. Motor unit potentiation of the quadriceps and gastrocnemius were found to be virtually unaffected by speed and incline variations as respects significant difference between the changes in these variables, although trends may exist between increased speed/incline and increased EMG activity, but this was not analyzed in this study.

Inclines may have no effect in a 0 g environment, however, restraint systems to treadmill exercises can be set such that forward inclinations in body lean can be made to produce resultant forces to bring about greater ground reaction forces. Increased G forces imposed by treadmill parameters imply greater transient forces to long bones involved in locomotion and therefore greater benefit to the structural integrity of bone tissue.

The study also revealed that physiological workload equivalencies, as set by MET requirements, are not paralleled by biomechanical and neuromuscular factors in actual exercise on the treadmill. Hence exercise countermeasures in space flight, where muscle, bone and neuromuscular considerations are concerned necessitate monitoring of the effect of the exercise regimens and modalities according to the response of the human body to the imposed demands of exercise by EMG and force transduction.

As this serves as a pilot study, the results lead to the following recommendations:

1. A basic 1 g comparison between the effects on motor unit potentiation (MUP) and ground reaction force (GRF) in treadmill and bicycle ergometer exercise.
2. The effects of prolonged treadmill running on bone density versus the same in bicycle ergometer exercise.
3. An analysis of the muscular torques and transarticular forces about the knee and ankle in treadmill running by means of optoelectrical data collection and anthropomorphic modeling.
4. The effects of treadmill running at various body inclinations and speeds on MUP and GRF in 0 gravity.
5. The differences between total body standing resistive exercises and treadmill running on MUP, GRF, muscle torques, transarticular reaction force and bone density.
6. The effects on MUP and GRF in decline treadmill running: an analysis of eccentric contractions in locomotion.
7. The effects of running shoes on MUP and GRF at different speeds and inclines in treadmill running.
8. The relationships between increasing speed/incline on MUP and GRF.

9. The relationships between MUP and physiological measures of energy expenditure (ie.  $\dot{V}O_2$  and kcal).

ORIGINAL PAGE IS  
OF POOR QUALITY

## REFERENCES

1. Cavangh, P.R., Biomechanical perspectives on locomotion in zero gravity, Workshop on Exercise Considerations in Long Duration Space Flight, NASA/JSC, July 1986.
2. Edgerton, R., Exercise issues related to the neuromuscular function and adaptation to microgravity, Workshop on Exercise Considerations in Long Duration Space Flight, NASA/JSC, July 1986.
3. Fremion, A., The effects of disuse on muscle, Workshop on Exercise Considerations in Long Duration Space Flight, NASA/JSC, July 1986.
4. Gazenko, O.G., Grigor'yev, A.I., and Notochin, Yu.V., Fluid-electrolyte homeostasis and weightlessness. From JPRS, USSR Report: Space Biology and Aerospace Medicine, 14(5):1-11, 1980.
5. Goode, A.W. and Rambaut, P.C., The skeleton in space, Nature, 317:204-05, 1985.
6. Guyton, A.C., Textbook of Medical Physiology, 7th edition, W.B. Saunders Co., Philadelphia, 1987.
7. Herbison, G.J. and Talbot, J.M., Muscular atrophy during space flight: Research needs and opportunities, The Physiologist, 28(6):520-26, 1985.
8. Komi, P.V., Physiological and biomechanical correlates of muscle function: Effects of muscle structure and stretch-shortening on force and speed. In R.L. Terjung (ed.), Exercise and Sport Science Review, 12:81-121, 1984.
9. Laffey, E.V., Nicogossian, A.E., Hoffler, G.W., Hurstra, W., and Baker, J., Spectral analysis of skeletal muscle changes resulting from 59 days of weightlessness in Skylab II (Report No. JSC 09996). NASA TM X-58171. Houston, TX: LBJ Space Center, 1975.
10. Laffey, E.V., Nicogossian, A.E., and Hoffler, G.W., Electromyographic analysis of the skeletal muscle changes arising from long and short periods of spaceflight weightlessness. Preprints of the 47th Aerospace Medical Association Meeting, Bal Harbour, FL, May 1976.
11. Moore, T.P., The history of inflight exercise in the United States manned space program, Workshop on Exercise Considerations in Long Duration Space Flight, NASA/JSC, June 1986.

ORIGINAL PAGE IS  
OF POOR QUALITY

12. Nicogossian A.E. and Parker, J.F. (eds.), Space Physiology and Medicine, NASA SP-447, 1982.

13. Stauber, W.T., A unique problem of muscle adaptation from weightlessness - The decelerator deficiency, Workshop on Exercise Considerations in Long Duration Space Flight, NASA/JSC, June 1986.

14. Thornton, W. and Rummel, J., Muscular deconditioning and its prevention in space flight. In R.S. Johnston and L.F. Dietlein (eds.), Biomedical Results from Skylab, NASA SP-377, 1977.

15. Thornton, W., Work, exercise and space flight II. Modification of adaptation by exercise, Workshop on Exercise Considerations in Long Duration Space Flight, NASA/JSC, June 1986.

16. Thornton, W., Work and exercise in space flight III. Exercise devices and protocols, Workshop on Exercise Considerations in Long Duration Space Flight, NASA/JSC, June 1986.

17. Volozhin, A.I., Didinko, I.Ye., and Stupakov, G.P., Chemical composition of mineral components of human vertebrae and calcaneus after hypokinesia. From JPRS, USSR Report: Space Biology and Aerospace Medicine, 15(1):60-3, 1981.

18. Vose, M.W., Review of roetgenographic bone demineralization studies of the Gemini space flights, Am J Roetgenol Radium Ther Nucl Med, 121:1-4, 1974.

19. Whittle, M.W., Herron, R., and Cuzzi, J., Biostereometric analysis of body form. In R.S. Johnston and L.F. Dietlein (eds.), Biomedical Results from Skylab, NASA SP-377, 1977.

20. Yergov, A.D., Results of medical research during the 175-day flight of the third prime crew on the Salyut-6/Soyuz orbital complex (NASA TM-76450). Moscow: Academy of Sciences USSR, 1980.

AN ASSESSMENT OF THE MICROMETEORITIC COMPONENT  
IN THE MARTIAN SOIL

Final Report

NASA/ASEE Summer Faculty Fellowship Program--1988

Johnson Space Center

Prepared By: George J. Flynn, Ph.D.  
Academic Rank: Assistant Professor  
University & Department: State University of New York  
Department of Physics  
Plattsburgh, NY 12901

NASA/JSC

Directorate: Space and Life Sciences  
Division: Solar System Exploration Division  
Branch: Space Resources Utilization Office  
JSC Colleague: David S. McKay, Ph.D.  
Date Submitted: Sept. 1, 1988  
Contract Number: NGT 44-005-803

# ABSTRACT

Particles in the mass range from  $10^{-7}$  to  $10^{-3}$  grams contribute 80% of the total mass influx of meteoritic material in the  $10^{-13}$  to  $10^6$  gram mass range at Earth (Hughes, 1978). On Earth atmospheric entry all but the smallest particles in the  $10^{-7}$  to  $10^{-3}$  gram mass range, about 60 to 1200 micrometers in diameter, are heated sufficiently to melt or vaporize. Mars, because of its lower escape velocity and larger atmospheric scale height, is a much more favorable site for unmelted survival of micrometeorites on atmospheric deceleration. We calculate that a significant fraction of particles throughout the 60 to 1200 micrometer diameter range will survive atmospheric entry unmelted. Thus returned Mars soils may offer a resource for sampling micrometeorites in a size range uncollectable in unaltered form at Earth.

The addition of meteoritic material to the Mars soils should perturb their chemical composition, as has been detected on for soils on the Moon (Anders, et al., 1973). Using the measured mass influx at Earth and estimates of the Mars/Earth flux ratio, we estimate a mass influx at Mars of between 2,700 and 202,000 metric tons per year. If distributed uniformly into a soil with a mean planetary production rate of 1 meter per billion years, consistent with radar estimates of the soil depth overlaying a bouldered terrain in the Tharsis region (Christensen, 1986), we estimate a meteoritic concentration in the soil ranging from 2 to 58%. This range is consistent with chemical inferences that Mars soil at the Viking sites is fit by a mixture of 60% indigenous rock fragments and 40% meteoritic material (Clark and Baird, 1979).

## INTRODUCTION

A number of sources for dust on Mars have been proposed including: chemical and physical weathering of indigenous surface rocks, debris from impact cratering, volcanism, tectonism (Greely, 1986), and weathering of shock-activated meteoritic projectiles from the period of heavy bombardment (Boslough, 1988). However, the contribution from micrometeorites, particles sufficiently small to be slowed without melting on atmospheric deceleration, has only recently been considered (Flynn and McKay, 1988). Particles in the size range which would be expected to survive atmospheric entry on Mars without melting constitute a major fraction of the meteoritic material presently accreting onto the Earth. On Earth, the rate of production of new soil is sufficiently high that unmelted particles are difficult to extract from the soils, however the meteoritic component can be detected by enrichment in Ir (Kyte and Wasson, 1986), and the particles themselves can be recovered from the stratosphere (Brownlee, 1985) and polar ices (Zolensky, et al., 1988), where the terrestrial soil contamination is reduced. Partially melted particles can also be extracted from sea sediments by magnetic concentration (Brownlee, 1985). Although particles in this size range are destroyed on Lunar impact, the Lunar mare soils contain between 1 and 2 percent meteoritic material, inferred from chemical abundances of the volatile and siderophile elements which are abundant in primitive meteorites but rare in the mare rocks (Anders, et al., 1973).

Mars, because of its low surface gravity coupled with an atmosphere of sufficient density to decelerate incoming micrometeorites, is one of the most favorable sites in the solar system for the unaltered survival of micrometeorites on atmospheric entry. The magnitude of the micrometeorite component in the Martian soils will depend on the accretion rate of meteoritic material onto the planet, as well as the rate at which soil is being produced by indigenous processes on Mars. To assess the importance of the micrometeorite component in the soil on Mars, we have:

- 1) assessed the micrometeorite flux at Mars,
- 2) estimated the micrometeorite velocity distribution at Mars, and evaluated the survival probability for micrometeorites entering the Martian atmosphere,
- 3) considered the chemical perturbations to the indigenous soil produced by micrometeorite addition, and suggested measurements which a Mars Lander spacecraft could make to determine the fraction of meteoritic material in the soils, and,
- 4) assessed the possibilities for recovery of micrometeorites from the soils returned by a Mars Sample Return mission. and,

Four types of meteoric and/or meteoritic material should be found on Mars:

- 1) micrometeorites, many of which will survive atmospheric deceleration unmelted, which should fall relatively uniformly over the planet's surface,
- 2) ablation products from larger meteors and meteorites which ablate, break up and/or burn up in the Mars atmosphere,
- 3) debris from large, crater forming objects, which, by analogy to terrestrial and lunar impact events, will be concentrated in the crater ejecta blankets (except for rare, large events, such as the proposed C-T event on earth, which can distribute debris on a planetary scale), and
- 4) debris from the early, intense bombardment, which, in many areas of the planet, may now be incorporated into rocks by geologic processes subsequent to the intense bombardment era.

## FLUX AT MARS

To estimate the extent of meteoritic addition to indigenous Martian soil, the meteoritic flux at Mars must be known. The meteoritic flux

measured at Earth provides a starting point to estimate the flux at Mars. For particles in the mass range from  $10^{-13}$  grams to  $10^6$  grams, the size-frequency distribution at Earth has been estimated from satellite, radar meteor and visual meteor observations (Hughes, 1978). Fifty percent of the total mass influx is in the narrow mass range from  $10^{-6}$  grams to  $10^{-4}$  grams (Hughes, 1978), and eighty percent is in the range from  $10^{-7}$  to  $10^{-3}$  grams. The Earth mass influx per decade of mass is shown in Table I. If the size-frequency distribution of meteoritic material at Mars is similar to that at Earth, only 0.53 au away, then an assessment of the Mars flux for material in the  $10^{-7}$  gram to  $10^{-3}$  gram mass range will provide a good estimate of the overall influx of meteoritic material on Mars.

On entering the Earth's atmosphere, most particles in the  $10^{-7}$  gram to  $10^{-3}$  gram mass range are volatilized, producing ionized trails detectable by radar, while some at the lowest end of this mass range are decelerated sufficiently slowly to survive entry unmelted (Brownlee, 1985). Veriani (1973) obtained the density distribution of over 5,000 radar meteors in the  $10^{-5}$  gram to  $10^{-3}$  gram mass range. More than 50% of his meteors fell in the density range of  $10^{-0.6}$  gm/cm<sup>3</sup> to  $10^{+0.4}$  gm/cm<sup>3</sup>, with a mean of the density distribution at 0.8 gm/cm<sup>3</sup>. Taking 1 gm/cm<sup>3</sup> as a representative density for the particles, the  $10^{-7}$  gram to  $10^{-3}$  gram mass range corresponds to particles from 60 to 1200 micrometers in diameter. Particles near the lower end of this mass range, up to about 100 micrometers in diameter, have been collected from the Earth's stratosphere, after atmospheric deceleration (Brownlee, 1985). The properties of these micrometeorites, collected in the NASA Cosmic Dust Sampling Program, have been reviewed by Brownlee (1985).

Most of the material in the 100 to 3000 micrometer size range incident on the top of the Earth's atmosphere would be expected to melt or volatilize during atmospheric deceleration (Brownlee, 1985). However Mars is a much more favorable site for particle survival on atmospheric deceleration since:

- 1) the scale height of the atmosphere on Mars is larger than on Earth, resulting in a longer deceleration time (and lower peak temperature on Mars than on Earth). and

- 2) the lower planetary mass and surface gravity give rise to less gravitational infall acceleration, thus particles with similar in-space velocities prior to planetary encounter will enter the atmosphere on Mars with lower velocity than on Earth.

The current mass influx at Earth in the  $10^{-13}$  gram to  $10^6$  gram mass range is estimated, from satellite, radar meteor, and visual meteor observations, at 16,100 tons per year, with 13,000 tons per year in the  $10^{-7}$  gram to  $10^{-3}$  gram mass range (Hughes, 1978). The measurement of Ir, which is believed to be a signature of extraterrestrial material, in atmospheric particulate samples at the South Pole suggest a current flux of extraterrestrial material of 11,000 tons per year (Tuncel and Zoller, 1987), consistent with the mass influx obtained by Hughes (1978).

Measurements of the long term meteoritic influx, by Ir concentrations in Pacific ocean sediments, have given values higher than the current flux. Kyte and Wasson (1986) infer a relatively constant mass influx (except for sharp spikes corresponding to major impact events) of 78,000 tons per year over the past 67 million years from Ir in Pacific sediments. This is consistent with an earlier value of 110,000 tons per year derived from Pacific sediments by Shedlowsky and Paisley (1966). We will perform our calculations using both the current Hughes (1978) mass influx at Earth, and the Kyte and Wasson (1986) long-term mass influx, which is a factor of five higher than the Hughes value. Literature values for the extraterrestrial mass influx at Earth, which are tabulated in Tuncel and Zoller (1987), exceed even this range. Since the Ir measurements provide no information on the incoming size distribution of the particles, we will assume the current size distribution from Hughes (1978) is also representative of the size distribution of the meteoritic material contributing to the long-term flux determined by Ir concentration (Kyte and Wasson, 1986). Some confirmation

of this assumption is provided by the analysis of the size versus frequency distribution of microcraters on exposed lunar rock surfaces returned by the Apollo missions. For particles larger than  $10^{-9}$  grams, where the effects of secondary impacts are no longer believed to be significant, the shape of the long-term lunar microcrater size versus frequency distribution is generally consistent with the mass distribution inferred from current satellite and meteor measurements (Grun, et. al, 1985).

To extrapolate the mass influx at Earth to the corresponding Mars value requires an estimate of the ratio of the Mars flux to the Earth flux. This ratio depends on the type of orbital evolution experienced by the particles. The orbits of large meteorites are perturbed principally by gravitational interactions with the planets. However, for small particles solar radiation pressure forces cause significant orbital perturbations on time scales comparable to or shorter than the gravitational perturbation time scale (Dohnanyi, 1978).

#### LARGE OBJECTS

For large, crater producing, objects whose orbits are dominated by gravitational perturbations, the relative crater production rates on the terrestrial planets and the Moon have been assessed to establish chronologies. Anders and Arnold (1965) have estimated the meteoritic input on Mars to be 25 times the Lunar value, while Soderblom, et al. (1974) have estimated the input of meteoritic material on Mars to be only twice that on the Moon. Shoemaker (1977) estimated that the ratio of impact rates of bodies to absolute visual magnitude 18 on Mars and Earth is 2.6. When corrected for differences in impact velocity, and scaled for the planetary surface gravities, Shoemaker (1977) arrived at a cratering rate ratio of 1.6. Hartmann et al. (1981) have reviewed the planetary cratering rate estimates, and they adopt a best value for the crater production rate on Mars of 1.3 that for Earth. Hartmann et al. (1981) indicate there are factor of 2 uncertainties in the cratering rate ratio, however, since the proportions of objects in various types of Earth and Mars crossing orbits are not well established. Since the Hartmann, et al. (1981) best value of the cratering ratio is consistent with the Shoemaker (1977) value, we will adopt the Shoemaker (1977) impact rate ratio of 2.6 as indicative of the ratio of the mass influx for large, gravitationally perturbed, objects at Mars and Earth.

#### SMALL OBJECTS

Objects whose orbits are dominated by P-R drag may, however, have a different ratio of the Mars to Earth flux. For particles up to 100 micrometers in diameter, the dominant radiation effect is a drag force, Poynting-Robertson (P-R) drag (Dohnanyi, 1978). For objects starting from circular orbits, P-R drag causes them to spiral into the sun. If the initial orbit of the particle is elliptical, P-R drag cause a rapid decrease in the perihelion and a slower decrease in the aphelion, so that the ellipticity of the orbit decreases as the particle falls into the sun.

The initial orbits of the dust particles will be determined by the orbits of the parent objects. The Infrared Astronomy Satellite (IRAS) detected two major types of solar system dust sources: the main belt asteroids (Low, et al., 1984), and comets (Sykes, et al., 1986). The dust bands detected in the main asteroid belt are thought to have been produced by low velocity collisions between main-belt asteroids (Sykes and Greenberg, 1986). Small particles produced in such collisions would spiral in towards the sun under P-R drag. They would pass Mars, providing an opportunity for Mars collection, and later pass Earth. Thus the Earth flux could be used to estimate the Mars flux for particles from this source region.

In the case of the particles emitted by comets, those detected by Sykes, et al. (1986) all had aphelia outside the orbit of Mars, and thus their orbits will also evolve, under P-R drag, through Mars collection and

subsequent Earth collection opportunities. Only if significant dust sources existed between the Earth and Mars, would there be a category of particles which, under P-R drag would be collectable at Earth but not at Mars. No such sources were reported in the IRAS survey.

To assess the ratio of the particle flux at Mars to that at Earth, the velocity distribution of the particles must be known at both planets. As shown by Opik (1951), the effective planetary capture cross-section,  $\sigma$ , varies with the velocity of the incident particle,  $v_p$ , as:

$$\sigma = \pi R_p^2 \cdot (1 + v_e^2/v_p^2) \quad (\text{Equation 1})$$

where  $v_e$  is the planetary escape velocity, and  $R_p$  is the planetary radius. The velocity distribution at Earth for radar meteors, particles in the  $10^{-6}$  to  $10^{-2}$  gram mass range, has been determined by Southworth and Sekanina (1973) for a set of over 14,000 radar meteors. We have used the Zook (1975) approximation to the Southworth and Sekanina (1973) atmospheric entry velocity distribution at Earth:

$$F(v) = 3.822 \times e^{-0.2468v} \quad (\text{Equation 2})$$

where  $F(v)$  is the fraction of the particles having a velocity  $v$ . This distribution is shown in Figure 1. We then followed the same procedures used by Morgan, et al. (1988), who calculated the meteoritic velocity distribution at Mercury, to calculate the velocity distribution at Mars. First, the Earth entry velocity distribution was corrected to an in-space distribution at 1 au by removing the near-earth gravitational focusing in each velocity increment, and removing the effect of earth infall acceleration. Next, the resulting velocity distribution was transformed to 1.53 au. At this time, the difference in flux at 1.53 au was also accounted for, taking the flux fall off with increasing heliocentric distance to vary as  $r^{-1.5}$ , as determined from zodiacal light observations (Hanner, et al., 1976; Schuerman, 1980). We then transformed the space velocity distribution at 1.53 au to a Mars atmospheric entry velocity distribution taking into account both the Mars gravitational focusing effect and the gravitational infall acceleration. The resulting atmospheric entry velocity distribution is shown in Figure 1. The ratio of the area under the curve of Earth entry velocity, which was normalized to 1.0, to the area under the Mars velocity distribution curve (0.57) is the flux ratio. This flux ratio must be multiplied by the planetary cross-sectional area ratio, equal to 0.29, to obtain the mass influx ratio. Thus, we estimate that, for a population dominated by P-R drag, the ratio of the mass influx at Mars to that at Earth would be 0.17.

The time for an equivalent amount of orbital evolution under P-R drag increases linearly with particle diameter if the density and reflectivity remain constant. Thus, as particles become larger they are more likely to be gravitationally perturbed from their orbits or destroyed by catastrophic collisions before P-R drag produces a significant change in the orbit. Estimates of the catastrophic collision lifetimes for particles in the main asteroid belt (Dohnanyi, 1978) and in space at 1 au (Grun, et al., 1985) both suggest that particles larger than about 100 micrometers in diameter will be destroyed by collisions before their orbits are perturbed into the sun. Uncertainties in the flux of particles producing collisional destruction might change this size limit.

#### FLUX ASSESSMENT

Thus particles larger than 100 micrometers diameter, which contribute 75% of the mass influx at earth, are expected to be only slightly larger than the size range whose orbital evolution is dominated by P-R drag. The actual ratio of the Mars to Earth flux is thus likely to be somewhere between that which we have calculated for P-R drag dominated particles, and

the value obtained by Shoemaker (1977) for larger objects whose orbital evolution is dominated by gravitational perturbations.

In assessing the meteoritic infall on Mars, we will take as a lower estimate the a value of 2700 tons per year obtained using the Hughes (1978) earth flux coupled with our P-R drag dominated flux ratio. An upper estimate of the Mars infall, of 202,000 tons per year, was obtained using the Kyte and Wasson (1986) terrestrial flux and the (Shoemaker, 1977) flux ratio. The large range in values reflects the uncertainties in the Earth flux and the Mars/Earth flux ratio.

If distributed uniformly over the planet, the corresponding mass accretion rates range from 18 grams per square meter per million years to 1350 grams per square meter per million years. For density  $1 \text{ gm/cm}^3$  material, these correspond to the addition of between 1.8 cm/billion years and 1.35 meters/billion years of meteoritic material to the Martian surface. Even the lower estimate would be expected to produce some detectable perturbations in the soil composition, unless the indigenous materials are very similar to meteoritic composition or the soil production rate far exceeds that on Earth.

Of this infalling mass, about 80% would be expected to be in the  $10^{-7}$  gram to  $10^{-3}$  gram mass range.

#### MICROMETEORITE SURVIVAL

Deceleration in the atmosphere of the Earth heats most of the particles in the  $10^{-6}$  gram to  $10^{-3}$  gram mass range above their melting temperature. Some particles in this mass range are recovered as melted spherules from the ocean bottom (Brownlee, 1985) and from pools in Greenland (Maurette, et al., 1988). However no extraterrestrial particle larger than approximately 50 micrometers in diameter has been recovered unmelted, and intact in the NASA Cosmic Dust Collection Program (Zolensky, pers. comm.).

Brownlee (1985) estimates the melting temperature for micrometeoritic material at about 1600 K. The distribution of peak temperatures reached by a micrometeorite on atmospheric entry can be predicated using an entry heating model developed by Whipple (1950) and extended by Fraundorf (1980). This temperature distribution is generally expressed as the fraction of particles heated above any given temperature. We show in Figure 2, the calculated fraction of incident particles of diameter 10 microns, 100 microns, and 1000 microns, heated above temperatures from 300 K to 2000 K on Earth atmospheric entry. The calculations use the Fraundorf (1980) equations, and assume the Southworth and Sekanina (1973) entry velocity distribution, a particle density of  $1 \text{ gm/cm}^3$ , and a value of 1 for the parameters characterizing drag, kinetic energy transfer, and emissivity. The reasons for this latter assumption are examined in Flynn (1988).

The general validity of the Fraundorf (1980) entry heating model can be verified by noting that, if 1600 K is taken as the critical temperature for unmelted survival, essentially all 10 micron particles, about half of the 100 micron particles, and essentially none of the 1000 micron particles would be expected to survive Earth atmospheric entry unmelted. This is consistent with the size cutoff on unmelted particles recovered from the stratosphere, and the observation that larger meteoritic material recovered from sediments is mostly melted.

We have applied the Fraundorf (1980) entry heating model to the Mars atmosphere. Within 30 km of the surface, the atmospheric scale height measured by Viking 1 was 11.70 km, and by Viking 2 was 11.36 km (Seif and Kirk, 1977). However spacecraft deceleration data for the atmosphere from 30 km to 120 km (Seif and Kirk, 1977) give a scale height in this region of 7.9 km. Since particles in the size range of interest reach their maximum deceleration in the upper region, we have used a scale height of 7.9 km. The velocity distribution which we inferred for Mars entry (Figure 1) was used, and the particle size, density, and interaction parameters were the

same as for the Earth atmospheric entry calculations. These results are also shown in Figure 2.

Unlike the Earth case, almost 90% of the 100 micron diameter particles would be expected to survive Mars atmospheric entry without melting. For the 1000 micron diameter particles, 30% would be expected to survive atmospheric entry unmelted. Micrometeorites in this size range which have survived atmospheric entry are rare in the Earth collections. Since these larger particles from Mars orbital distance are likely to sample different sources than the smaller micrometeorites collected at 1 a.u. in the cosmic dust sampling program on Earth (Flynn, 1988; Zook and McKay 1986), returned Mars soils may provide a unique resource for micrometeorite analysis.

We have used our calculated Mars atmospheric entry velocity distribution, and the Fraundorf (1980) entry heating model, to plot the predicted fraction of incoming micrometeorites not heated above temperature  $T$  on Mars atmospheric deceleration for temperatures ranging from 700 K to 1900 K. These results are shown for particles diameters from 10 to 1000 micrometers in Figure 3. We have used this survival fraction (those not heated above 1600 K at each size), coupled with the size-frequency distribution of micrometeorites from Hughes (1978), and our high and low estimates of the total mass influx at Mars to calculate the micrometeorite addition rate (particles per square meter of surface per year) to the Martian soils. Since each decade of mass spans only a factor of 2.1 in particle diameter, and since the particle abundance is a rapidly decreasing function of mass, all particles within each mass decade are taken to be at the smallest diameter in that decade. Particles are assumed to have a density of  $1 \text{ gm/cm}^3$ , consistent with the range of 0.7 to  $2.2 \text{ gm/cm}^3$  measured for the smaller micrometeorites recovered from the Earth's stratosphere (Fraundorf, et al., 1982). These results of these calculations are reported in Table II. The expected abundance of micrometeorites in a returned Mars soil sample could be estimated from these results if the soil production rate on Mars were known.

Estimates of the thickness of the Martian regolith vary widely, from only twice as thick as the lunar regolith (Soderbloom, et al., 1974) to as deep as 2 km (Fanale, 1976). However, much of the planetary regolith was very likely generated during the intense bombardment era. Depending on the mixing depth, the present meteoritic infall may or may not be mixed into the soil of that early regolith.

Various physical properties of the Martian surface can be used to constrain the thickness of the current dust deposits. The low thermal inertia of the deposits requires a minimum thickness of order 0.1 meters (Harmon, et al., 1982). Harmon, et al. (1982) also suggest that the presence of exposed rocks and the degree of visible mantling indicate the dust thickness is less than 5 meters. Dual-polarization radar measurements in the Tharsis region indicate a rough texture, which suggests that a relatively thin dust layer covers near-surface rocks. Based on radar penetration properties, Christensen (1986) estimates a dust mantle thickness of only 1 to 2 meters. Arvidson (1986) suggests that most of the sedimentary debris on Mars was produced relatively early, perhaps in the first billion years. In more recent times, the preservation of a large number of pristine-looking, small bowl shaped craters at the Viking 1 lander site suggests a rate of rock breakdown and removal of only meters per billions of years (Arvidson, 1986).

Thus, while the regolith itself could be quite deep (Fanale, 1976), much of it is likely to have been produced during the intense bombardment of the planet. The surface soil into which the last billion years of meteoritic material may be concentrated, could be only a few meters deep.

#### DIRECT COLLECTION

We calculate the expected concentration of micrometeorites in a soil whose planetary average production rate is 1 meter per billion years. The concentrations obtained can then be scaled to other assumed soil production

rates by multiplying by (1 meter/billion years)/(production rate). Table III shows the number of unmelted micrometeorites in each size range expected in an average 10 gram Mars soil sample. These results suggest that returned Mars samples may offer a new resource for the study of micrometeorites. These micrometeorites may sample a different parent population than the smaller particles recovered from the Earth's stratosphere. The survival lifetime of micrometeorites deposited in the harsh Martian environment is unknown, however on Earth millimeter size fragments of meteoritic material, both unmelted and melted, have been recovered from late Pliocene deep sea sediments (Brownlee, et al., 1982), demonstrating that survival is possible on Earth for several million years.

The soils would also be expected to contain melted micrometeorites in the larger sizes. On Earth, such particles can easily be extracted from the ocean bottom by magnetic separation, due to the formation of magnetic minerals on atmospheric entry (Brownlee, 1985). Depending on the abundance of indigenous magnetic minerals in the Mars soil, a magnetic separation may permit recovery of the melted particles to determine if the meteoritic concentration is sufficient for search and extraction of unmelted particles.

#### COLLECTING SITES FOR MICROMETEORITES

Martian surface processes (weathering and wind erosion, transport, and deposition) may fractionate the dust by size, density or composition providing regions of increased local concentration, suggesting even more suitable sites for micrometeorite sampling than the average soil. These sites may include placer catch basins or lag surfaces which may accumulate high density micrometeorites or their derived and altered minerals. Conversely, low density micrometeorites may be wind segregated along with finer Martian dust and may constitute a relatively coarse-grained component of that dust at its deposition sites. By analogy with Antarctica, meteorites of all size ranges may be relatively concentrated in Martian polar regions, although the concentration mechanisms may be different.

#### MICROMETEORITES AS A TOOL

Assuming that micrometeorites could be identified in returned soil samples, this addition of micrometeorite material to the uppermost Martian regolith at a constant rate could conceivably provide a powerful tool for tracking rates of erosion, deposition, and weathering. On Mars Sample Return missions, an attempt should be made to collect soils from different geologic sites (catch basins, lag surfaces, flat high plains, valley bottoms, etc.) so as to provide a variety of soils of different sedimentary environments. One of the important differences among these environments might be the proportion of petrographically or chemically identifiable micrometeorites mixed into the soil.

#### CHEMICAL SIGNATURES

The possibility that detectable micrometeorites and their remains can be found in the Martian soils depends on the relative rates of infall, weathering and alteration, transportation, and mixing. These rates are not yet known reliably enough to allow us to predict with certainty whether identifiable micrometeorites will be found. While they may be relatively quickly destroyed by Martian weathering, the chemical signatures, particularly siderophiles and volatiles, may persist in the soils, as they have in the lunar regolith (Anders, et al., 1973), and in Earth sediments where Ir anomalies are detected (Kyte and Wasson, 1986). All of the meteoritic material collected by the planet, whether it reaches the surface unmelted, or if it melts or vaporizes on atmospheric entry, should eventually be added to the soils. We have estimated a mass influx on Mars ranging from 2,700 to 202,000 tons per year. If spread uniformly over

the surface of the planet, this corresponds to an addition rate from  $1.8 \times 10^{-5} \text{ gm/m}^2 \cdot \text{year}$  to  $1.4 \times 10^{-3} \text{ gm/m}^2 \cdot \text{year}$ . For material of density  $1 \text{ gm/cm}^3$ , this corresponds to an accretion rate which ranges from 1.8 centimeters per billion years to 1.4 meters per billion years.

If the soil production rate on Mars is of order 1 meter per billion years, these meteoritic accretion rates would give rise to meteoritic concentrations ranging from 2% to 58% in the average Martian soil. Boslough (1988) and Clark and Baird (1979), applying the subtraction method to the Viking chemical data, suggest the Mars regolith can be fit by a mixture of 40% CI meteorite and 60% planetary rock fragments. Boslough (1988) suggests the meteoritic component is ancient. But it could equally well be the micrometeorite component, which dominates the ancient component in lunar mare soils. In the lunar case the composition of the non-indigenous material was taken as the residual after subtracting rock composition from soil composition. Two meteoritic components were detected. In mature soils the residual has a trace element composition consistent with the addition of 1.5% CI meteoritic material, attributed to the long term micrometeorite infall (Anders et al., 1973). Less mature highland soils also show a second component, characterized by fractionated siderophile content and low volatiles, attributed to ancient bombardment (Anders et al., 1973).

The observations of Boslough (1988) and Clark and Baird (1979) are consistent with the Mars soil at the Viking sites containing a substantial meteoritic component. The meteoritic concentration they infer is consistent with the range of concentrations we calculate from micrometeorite influx, provided the rate of production of soil on Mars is no more than a few meters per billion years. The possibility exists that the Martian soils contain a substantial fraction of meteoritic material. On Earth and on the Moon, the Ni/Fe ratio and the Ir abundance have proven to be diagnostic indicators of the meteoritic component, since Ir and Ni are enriched in CI meteorites but depleted in crustal materials. Direct spacecraft measurement of the Ni and/or Ir abundances in the Mars regolith should help determine the meteoritic content of the soil.

#### CONSTRAINTS FROM VIKING MEASUREMENTS

The major meteoritic component in the Lunar mare soils is similar, in siderophile and volatile element composition, to the CI carbonaceous chondrite meteorites. These meteorites contain an average of 3 to 5% carbon (Wasson, 1974), some in the form of organic matter. The cosmic dust particles collected from the Earth's stratosphere contain carbon at or above CI concentrations (Blanford, et al., 1988), some of which may be in the form of polycyclic aromatic hydrocarbons (Allamandola, et al., 1987). If the Martian soil contained a substantial abundance of unmodified CI carbonaceous chondrite material, the Viking gas chromatograph mass spectrometer should have detected its presence.

Biemann, et al. (1977) found no detectable organic material in four Martian samples, one surface and one subsurface at each of the two Viking sites. Using the laboratory version of the Viking gas chromatograph mass spectrometer, they detect naphthalene at a level of about 1 ppm in CI carbonaceous chondrite samples (Biemann, et al., 1977). They report a detection limit of 0.5 ppb for naphthalene at the Viking 1 site and 0.015 ppb for Viking 2 (Biemann, et al., 1977). If the concentration of naphthalene were the same in the infalling meteoritic material as in their carbonaceous chondrite sample, then the corresponding upper limits on the meteoritic mass fraction in the analyzed Viking samples would be 0.05% at the Viking 1 site, and 0.0015% at the Viking 2 site.

The organic content of the micrometeorites in the  $10^{-6}$  to  $10^{-2}$  gram mass range has never been established, since the particles in this mass range collected on Earth are melted on atmospheric entry. Furthermore, as pointed out by Banin (1988), the high redox potential of the Martian soil may have caused the decomposition of any organic matter from meteoritic infall.

## MARTIAN AGGLUTINATES AND SOIL MATURITY

If, as we calculate, micrometeorites are all slowed down by the Martian atmosphere, and assuming that most lunar agglutinates are made by micrometeorite impacts, no analogous Martian agglutinates would be expected (unless there were an era in which the atmosphere was considerably less dense than at present). However, many types of impact glasses would be expected from larger impacts, and some of these glasses may resemble lunar agglutinates in some respects.

Gault and Baldwin (1970) have estimated a minimum impact crater size of 50 meters, taking into account fragmentation and ablation of the incoming projectiles as well as atmospheric deceleration. The smallest craters noted in Viking orbiter images are about 100 meters in diameter (Blasius, 1976), but smaller craters beyond the resolution limit of the photographs may still be present. Dycus (1969) predicts that projectiles as small as 10 gm would still form craters. However, craters too small to be seen from the orbiter are not apparent in Viking lander images. Impact gardening associated with the 50 meter and larger craters predicted by Gault and Baldwin (1970) would determine regolith turnover rates and cause comminution of rocks into soils. The addition of micrometeorites would affect the petrology and chemistry of Martian soil. Weathering and sedimentary processes on Mars would also process the regolith components. The overall effect would be to make an exceedingly complex regolith. A new maturation scale will be necessary for Martian regolith. This scale will have to include terms which reflect (1) impact reworking, (2) addition of micrometeorites, and (3) Martian surface weathering and alteration. For example, if concentration mechanisms can be factored out, the abundance of micrometeorites (identified petrographically or chemically) in a soil layer might be directly related to its near-surface exposure time in a manner analogous to the abundance of agglutinates in lunar soils. In addition to soil evolution through maturation, physical mixing of soils of differing maturities should be common.

## CONCLUSIONS

Micrometeorites in the  $10^{-7}$  to  $10^{-3}$  gram mass range should be a major contributor to the meteoritic input on Mars. Unlike the Earth, where most particles above  $10^{-6}$  grams are melted or vaporized on atmospheric entry, a large fraction of these particles are expected to survive entry into the atmosphere of Mars unmelted. The addition of this meteoritic material to the Martian regolith could significantly perturb the chemical abundances in the soils, particularly the abundances of volatile and siderophile elements which are abundant in CI meteorites but depleted in crustal materials, and of noble gases (and possibly hydrogen) which are implanted in micrometeorites during space exposure and carried into the soils with the particles.

Uncertainties micrometeorite flux at Mars as well as the rate of production of soil through weathering processes on the planet give rise to large uncertainties in the meteoritic concentration in the Martian soils. However, our estimates are not incompatible with the suggestion by Boslough (1988) that the soils analyzed by Viking are 40% meteoritic.

The first returned soil samples from Mars should provide the opportunity for recovery and analysis of unaltered micrometeorites larger than any sampled on earth, assessment of the magnitude of the meteoritic component, and possibly an estimate of the rate of soil production on the planet. The larger micrometeorites which enter the atmosphere of Mars without melting may sample a different source population than is sampled by the smaller particles collected from the Earth's stratosphere.

# REFERENCES

- Allamandola, L.J., Sandford, S.A., and Wopenka, B. (1987). Interstellar Polycyclic Aromatic Hydrocarbons and Carbon in Interplanetary Dust Particles and Meteorites, *Science*, **237**, 36-38.
- Arvidson, R.E. (1986). On the Rate of Formation of Sedimentary Debris on Mars, in *Dust on Mars II*, LPI Tech Report 86-08, 8.
- Anders, E. and Arnold, J. (1963). Age of Craters on Mars, *Science*, **159**, 1404-1408.
- Anders, E., Ganapathy, R., Krashinsky, U., and Morgan, J.W. (1973). Meteoritic Material on the Moon, *The Moon*, **3**, 3-24.
- Bain, A. (1988) The Soils of Mars, *Mars Sample Return Science Workshop* (abstract volume), L1 11.
- Bismann, K., Oro, J., Toulmin, P. III, Orgel, L.E., Wier, A.O., Anderson, D.N., Simmonds, P., Flory, D., Oles, A.V., Bushneck, D.R., Miller, J.S., Lofgren, A.L. (1977). The Search for Organic Substances and Inorganic Volatile Compounds in the Surface of Mars, *Journ. Geophys. Res.*, **82**, 4641-46.
- Blenford, G.E., Thomas, E.L., and McKay, D.S. (1988). Microbeam Analysis of Four Chondritic Interplanetary Dust Particles for Major Elements, Carbon and Oxygen, *Meteoritics*, in press.
- Boelough, M.B. (1988). Evidence for Meteoritic Enrichment of the Martian Regolith, *LPSK XIX*, 120-121.
- Brownlee, D.E. (1985). Cosmic Dust: Collection and Research, *Ann. Rev. Earth Planet. Sci.*, **13**, 147-173.
- Brownlee, D.E., Kyte, F.T., and Wason, J.T. (1982). Unmelted Meteoritic Material Discovered the Late Pleistocene in Antarctica, *LPSK XIII*, 60-70.
- Christensen, P.R. (1986). Dust Deposition and Erosion on Mars: Cyclic Development of Regional Deposits, in *Dust on Mars II*, LPI Tech. Report 86-08, 10-15.
- Clark, B.C., and Beir, A.K. (1979). Is the Martian Lithosphere Sulphur Rich?, *Journ. Geophys. Res.*, **84**, 8383-8403.
- Dohnanyi, J.S. (1978). Particle Dynamics, in *Cosmic Dust*, (ed., J.A.N. McDermott), Wiley, New York, 527-605.
- Famale, F.P. (1976) Martian Volatiles: Their Degassing History and Geochemical Fate, *ICARUS*, **28**, 179-202.
- Flynn, G. J. (1988). Atmospheric Entry Heating of Micrometeorites, *Proc. Lunar Planet. Sci. Conf.*, **19th**, in press.
- Flynn, G.J., and McKay, D.S. (1988). Meteorites on Mars, *Workshop on Mars Sample Return*, LPI Tech Report, in press.
- Fraundorfer, P. (1980). The Distribution of Temperature Maxima for Micrometeorites in the Earth Atmosphere Without Melting, *Geophys. Res. Lett.*, **12**, 763-768.
- Fraundorfer, P., Hints, C., Lowry, O., McKee, K.D., and Sandford, S.A. (1982). Determination of the Mass, Surface Density, and Volume Density of Individual Interplanetary Dust Particles, *LPSK XI*, 225-226.
- Greeley, R. (1986). Toward an Understanding of the Martian Dust Cycle, in *Dust on Mars II*, LPI Tech. Report 86-08, 28-31.
- Grün, E., Zook, R.A., Feghtig, E., and Giese, R.B. (1985). Collisional Balance of the Meteoritic Complex, *ICARUS*, **28**, 244-272.
- Hanner, M.S., Sparrow, J. G., Weinberg, J.L., and Beeson, D.E. (1976). XXXXXXXX, in *Interplanetary Dust and the Zodiacal Light* (eds., Elascor, H. and Feghtig, E.), Springer-Verlag, Heidelberg, 29-35.
- Harmon, J.K., Campbell, D.B., and Ostro, S.J. (1982). *ICARUS*, **22**, 171-187.
- Hartmann, W.K., Strom, R.G., Weidenschilling, S.J., Blasius, K.R., Morozov, A., Dence, M.R., Griev, R.A.F., Diaz, J., Chapman, C.R., Shoemaker, E.M., and Jones, K.L. (1981). Chronology of Planetary Volcanism by Comparative Studies of Planetary Craters, in *Basaltic Volcanism on the Terrestrial Planets*, Pergamon Press, New York, 1049-1127.
- Hughes, D.M. (1978). "Meteorite," in *Cosmic Dust* (ed., J.A.N. McDermott), Wiley, New York, 123-185.
- Kyte, F.T. and Wason, J.T. (1986). *Science*, **232**, 1225.
- Low, F.J., Beinman, D.A., Goutier, T.R., Gillette, F.C., Reichman, C.A., Fougere, G., Yon E., Amann, H.E., Boggess, H., Emerson, J.F., Sabing, R.J., Sousser, M.G., Souch, J., Brown-Robinson M., Soifer, B.T., Walker, R.G., and Weisellus, P.R. (1984). Infrared Cirrus: New Components of the Extended Infrared Emission, *Astrophys. J.*, **278**, L19-L22.
- Mouretto, M., Samer, C., Fourchet, M., and Brownlee, D.E. (1988). The "Blue Lake II" Expedition of July-August 1987 in Greenland, and the Search for rare Unmelted Extraterrestrial Particles, *LPSK XIX*, 744-745.
- Moore, H.J., Banton, R.E., Scott, R.F., Spitzer, C.R., and Shorthill, R.W. (1977). Surface Materials the Viking Landing Sites, *Journ. Geophys. Res.*, **82**, 4497-4523.
- Morgan, T.S., Zook, R.A., and Potter, A.K. (1988). Impact-Driven Supply of Sodium and Potassium to the Atmosphere of Mercury, *ICARUS*, in press.
- Morgan, J.W., Krashinsky, U., Ganapathy, R., and Anders, E. (1973). Luna 20 Soil: Abundance of 17 Trace Elements, *Geochim. Cosmochim. Acta*, **37**, 933-941.
- Opik, E.J. (1951) Collision Probabilities with the Planets and the Distribution of Interplanetary Matter, *Proc. R. Irish Acad.*, **51**, Sect. A, 165-190.
- Schwarm, D.W. (1988), in *Solid Particles in the Solar System*, (eds., Holliday, I., and McIntosh, B.A.), D. Reidel, Dordrecht, 71-74.
- Seif, A., and Kirk, D.B. (1977). Structure of the Atmosphere of Mars in Summer at Mid-Latitudes, *Journ. Geophys. Res.*, **82**, 4364-4378.
- Shoemaker, E.M. (1977). Astronomically Observable Crater-Forming Projectiles, in *Impact and Explosion Cratering* (eds. Roddy, D.J., Pepin, R.O., and Merrill, R.B.), Pergamon Press, New York, 617-628.
- Shoemaker, J.P., and Paisley, S. (1966). *Tellus*, **18**, 490-503.
- Soderblom, L.A., Camdit, C.D., West, R.A., Barrow, S.M., and Kradler, T.J. (1974). Martian Planet-Wide Crater Distribution Implications for Geologic History and Surface Processes, *ICARUS*, **22**, 239-263.
- Southworth, S.A., and Sekizawa, E. (1973). *Physical and Dynamical Studies of Meteors*, NASA CR 2216.
- Sykes, M.V., Lebedev, L.A., Banton, D.M., and Low, F. (1988). The Discovery of Dust Trails: the Orbits of Periodic Comets, *Science*, **222**, 1115-1117.
- Sykes, M.V., and Greenberg, R. (1986). The Formation and Origin of the IRAS Zodiacal Dust Belt as a Consequence of Single Collisions Between Asteroids, *ICARUS*, **31**, 31-68.
- Tuncel, G. and Zoller, W.H. (1967). Atmospheric Iridium at the South Pole on a Measure of the Meteoritic Component, *Tellus*, **19**, 703-705.
- Verniani, F. (1973). An Analysis of the Physical Parameters of 3750 Faint Radio Meteors, *J. Geophys. Res.*, **78**, 6420.
- Wason, J.T. (1974) *Meteorites: Classification and Properties*, Springer, New York.
- Whipple, F.L. (1956). The Theory of Micro-Meteorites. Part I: In an Isothermal Atmosphere, *Proc. Nat. Acad. Sci.*, **48**, No. 12, 687-693.
- Zolensky, M.E., Webb, S.J., and Thomas, K. (1988). The Search for Interplanetary Dust Particles from Preindustrial Age Antarctic Ice, *Proc. Eighteenth Lunar Planet. Sci. Conf.*, Cambridge Univ. Press, 599-605.
- Zook, R.A. (1975). The State of Meteoritic Material on the Moon, *Proc. Lunar Sci. Conf.*, **8th**, 1653-1672.

ORIGINAL PAGE IS  
OF POOR QUALITY

Table I  
Mass Influx at Earth\*

Mass Range (grams)	Particle Diameter (microns) <sup>+</sup>	Mass Influx (kg/year)	Fraction of Total Mass <sup>#</sup>
10 <sup>-7</sup> to 10 <sup>-6</sup>	58 to 124	2.0 x 10 <sup>6</sup>	.12
10 <sup>-6</sup> to 10 <sup>-5</sup>	124 to 268	4.2 x 10 <sup>6</sup>	.26
10 <sup>-5</sup> to 10 <sup>-4</sup>	268 to 576	4.2 x 10 <sup>6</sup>	.26
10 <sup>-4</sup> to 10 <sup>-3</sup>	576 to 1240	2.7 x 10 <sup>6</sup>	.17
10 <sup>-3</sup> to 10 <sup>-2</sup>	1240 to 2600	1.3 x 10 <sup>6</sup>	.08

\* Data, except for particle diameters, from Hughes (1978).

# Ratio of mass in this decade to total mass influx from 10<sup>-13</sup> to 10<sup>6</sup> grams.

+ Diameters calculated for spheres of density 1 gm/cm<sup>3</sup>.

Table II  
Particles Surviving Mars Atmospheric Entry Unmelted

Size Range (micrometers)	Upper Estimate (particles/m <sup>2</sup> ·year)	Lower Estimate (particles/m <sup>2</sup> ·year)
56 to 124	1.6 x 10 <sup>3</sup>	2.1 x 10 <sup>1</sup>
125 to 268	3.2 x 10 <sup>2</sup>	4.3
269 to 576	2.7 x 10 <sup>1</sup>	3.6 x 10 <sup>-1</sup>
577 to 1240	1.2	1.6 x 10 <sup>-2</sup>
1241 to 2600	3.1 x 10 <sup>-2</sup>	4.1 x 10 <sup>-4</sup>

Table III  
Number of Unmelted Micrometeorites Expected  
in 10 gram Average Soil Sample

Size Range (micrometers)	Upper Estimate	Lower Estimate
56 to 124	16,000	210
125 to 268	3,200	43
269 to 576	270	4
577 to 1240	12	< 1
1241 to 2600	3	< 1

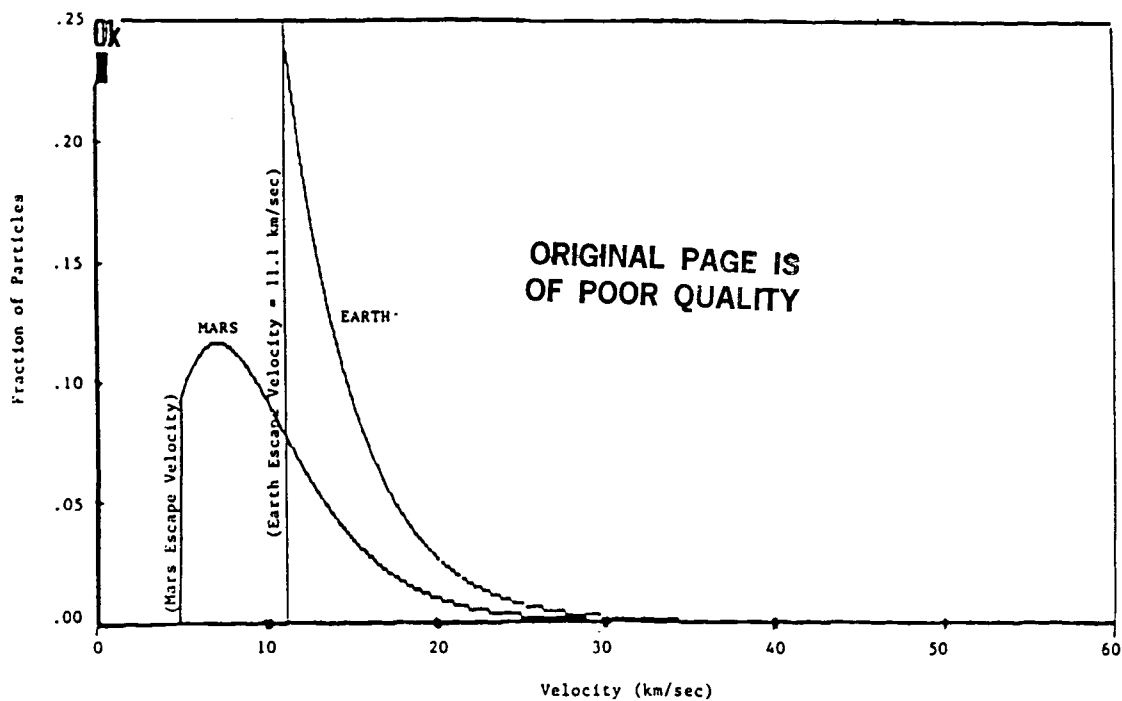


Figure 1.-- Atmospheric entry velocity distributions measured at Earth (Southworth and Sekanina, 1973), and calculated at Mars.

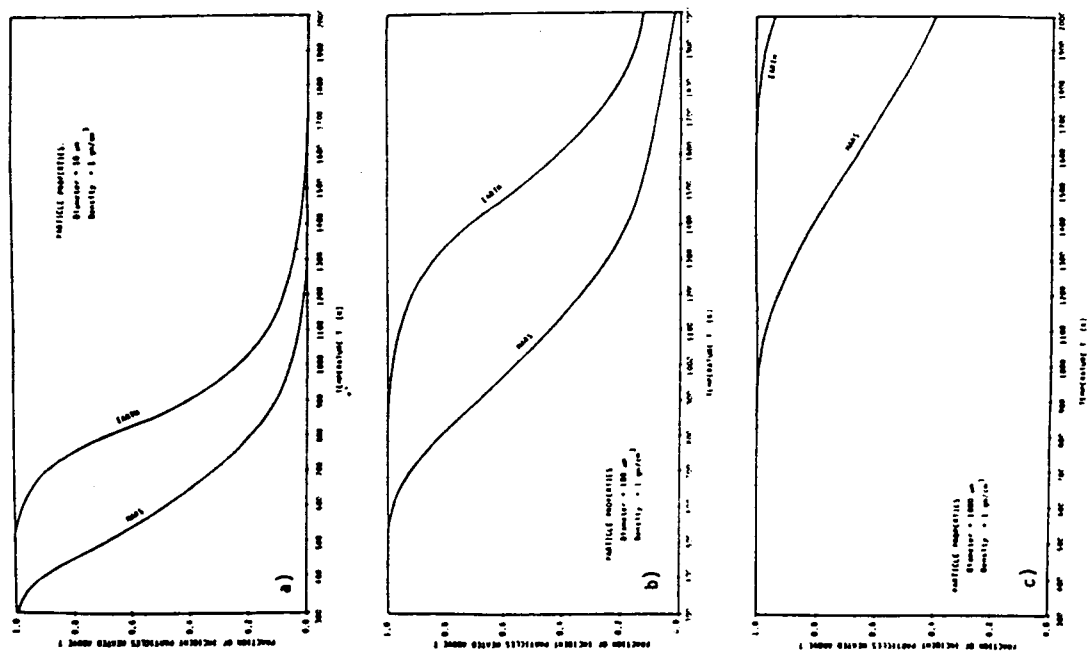


Figure 2.--Fraction of entering particles heated above temperature T on atmospheric deceleration for Earth and Mars for three particle diameters: a) 10 micrometers, b) 100 micrometers, and, c) 1000 micrometers.

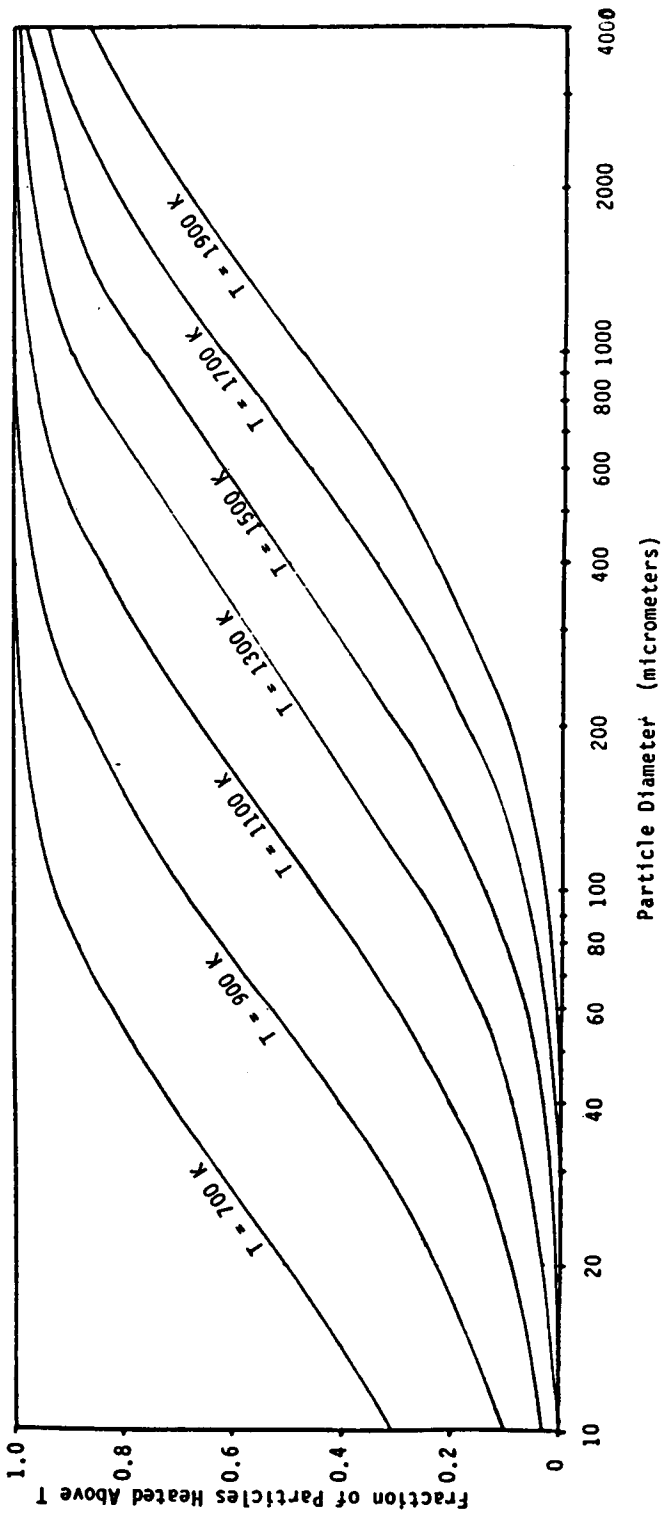


Figure 3.-- Mars atmospheric entry temperature distribution for particles having the Mars entry velocity distribution shown in Figure 2. Particles have a density of  $1 \text{ gm/cm}^3$ , and a thermal emissivity of 1.

N89 - 20071

EVALUATION OF AVAILABLE ANALYTICAL TECHNIQUES  
FOR MONITORING THE QUALITY OF  
SPACE STATION POTABLE WATER

Final Report

NASA/ASEE Summer Faculty Fellowship Program--1988

Johnson Space Center

Prepared By:	Richard D. Geer, Ph.D.
Academic Rank:	Associate Professor
University & Department:	Montana State University Department of Chemistry Bozeman, Montana 59717

NASA/JSC

Directorate:	Space and Life Sciences
Division:	Medical Sciences
Branch:	Biomedical Laboratories
JSC Colleague:	Richard L. Sauer, P.E.
Date Submitted:	September 30, 1988
Contract Number:	NGT 44-005-803

## ABSTRACT

To assure the quality of potable water on the Space Station a number of chemical and physical tests must be conducted routinely. After reviewing the requirements for potable water, both direct and indirect analytical methods are evaluated that could make the required tests and improvements are suggested be make them compatible with the Space Station operation.

Of the possible analytical procedures that might be carried out on the SS, the following seem to be required to meet the water quality criteria: turbidity and color from photometric measurements, a conductivity measurement, pH by an ion selective electrode, ion chromatography for inorganic ions with set MCLs, and a total organic carbon measurement. Also required are generic methods for total phenols (MCL 1 ppb), total organic halogen compounds (MCL 10 ppb), and organic acids (MCL 500 ppb). Electrochemical methods of controlled potential amperometry and microcoulometric titrimetry may meet this need. The other PW standards for organic alcohols (MCL 500 ppb) and TOC less nontoxicants (MCL 100 ppb), are broadly defined and can only be met with either an approximate surrogate measure (example, UV absorption of aromatic compounds for "toxic TOC"), or methods that separate and measure specific compounds in those classes that are likely to be found in PW. Finally the problem of analyzing for specific toxicants that are yet to be defined will either require specific sensors (electrode or optrode), or a separation and detection system.

Each of the separation techniques evaluated appear to have special areas of applications for the SS, though supercritical fluid chromatography with  $\text{CO}_2$  could probably replace gas chromatography while extending the range of analyses to nonvolatile heat sensitive compounds. Most forms of HPLC are not very compatible with the SS environment, but ion chromatography is needed for most anion and cation analyses (anodic stripping voltammetry could measure the heavy metal ions). It would be more suitable if it were converted to microbore column size. Capillary zone electrophoresis has the potential as a universal separation system, but it will require much more development in terms of buffer composition and sensitive micro volume detectors.

The most universal detector for the separation techniques is the high resolution mass spectrum detector, however its use on the SS is limited because of size, power and weight requirements associated with its high vacuum system. The need to have one on the SS seems justified only if it would be used for a variety of experiments and not just to test PW. The electrochemistry/conductivity detector appears to be the best compromise for a detector for HPLC or CZE.

Techniques are suggested for removing the interfering biocide  $\text{I}_2/\text{I}^-$  from water, and for carrying out chemical and electrochemical operations in microgravity where mixing and gas liquid separations are required.

## SPACE STATION POTABLE WATER

The potable water (PW) for use on the Space Station (SS) will come from the cabin air humidity condensate (HC) produced by the air revitalization system (ARS). This means the PW is in a direct recycle loop, a treatment process not previously implemented. To insure the PW remains safe for the duration of a 90 to 180 day mission, it needs to be continually monitored. This includes frequent or continuous testing of several important general or surrogate indicators that show the treatment system is operating properly. There are also a number of other specifications that the PW must continue to meet with regard to specific chemicals or classes of compounds. Table 1 gives the current criteria set for SS PW (1). The table does not include the microbiological standards, as they are not of concern to this report.

Unfortunately the majority of the organic carbon found in samples of HC from the two experiments that simulated the expected SS conditions (Space Lab and Bends) have not been correlated with specific organic compounds (2,3). Therefore, maximum contaminant levels (MCL) for possible toxic organic chemicals in PW can not be set until more specific information is obtained about the composition of the TOC in representative samples of HC from ground based tests. This, by itself, is a major analytical challenge and can influence the types of chemical analyses that are required on the SS.

This report will attempt to evaluate the kinds of on-board analytical procedure that will be required and also look at alternate methods that accomplish the same ends and/or add valuable data for determining PW quality. Modifications of procedures needed to meet the required sensitivity and potential interferences will be suggested along with ideas for efficient operation within the limitations of the SS environment. First the direct analytical methods which allow continuous on-line monitoring of the PW will be discussed, then the indirect methods which require discrete samples will be considered. These later methods require chemical reaction or separation and are the more sophisticated and complex analytical techniques.

## SPACE STATION OPERATIONAL LIMITATIONS

There are four rather severe limitations on the analytical chemistry procedures that can be used on the SS.

1. The biggest limitation is microgravity which does not allow the automatic phase separations of immiscible fluids. That is, gases and liquids do not separate without an external applied force. They can be separated by diffusion, but that is slow and inefficient. However, in zero G convection does not occur due to small density differences, this can be helpful for processes like electrochemistry or capillary zone electrophoresis (CZE) that will be discussed later.

TABLE 1.- SPACE STATION POTABLE WATER QUALITY REQUIREMENTS, MAXIMUM CONTAMINANT LEVELS (1).

Physical Parameter Limits

Total Solids (mg/l)	100	pH	6.0-8.0
Conductivity ( $\mu$ mhos/cm)	-	Turbidity (NTU)	1
Color, true (Pt/Co units)	15	Dissolved Gas (free at 35 C)	None
Taste & Odor (TTN/TON)	3	Free Gas (@ STP)	None

Inorganics mg/l (ppm)

Ammonia	0.5	Cyanide	0.2	Mercury	0.002
Arsenic	0.01	Fluoride	1.0	Nickel	0.05
Barium	1.0	Iodine		Nitrate	10.
Cadmium	0.01	Total	15.	Potassium	340.
Calcium	30.	Iron	0.3	Selenium	0.01
Chloride	250.	Lead	0.05	Silver	0.05
Chromium	0.05	Manganese	0.05	Sulfide	0.05
Copper	1.0	Magnesium	50.	Sulfate	250.
				Zinc	5.0

Organics	$\mu$ g/l (ppb)	Aesthetics	(mg/l)
** Total Organic Carbon (TOC)	500.	Cations	30
TOC (Less Nontoxicants)	100.	Anions	30
Generic Classes		CO <sub>2</sub>	15
Acids	500.		
Alcohols	500.		
Halogenated Hydrocarbons	10.		
Phenols	1.		
** Specific Toxicants	TBD		

\*\* Note, less than half the TOC from Space Lab humidity condensate samples have been associated with specific organic compounds.

2. Mission personnel will not be specifically trained in chemistry and their time will be limited for chemical or instrumental manipulations. This requires careful consideration of analytical techniques in terms of in-flight serviceability, frequency of instrument reconfiguration, recalibration or component failure, and consumable servicing.

3. Due to health and weight considerations, limits will be set on both the amount and toxic nature of consumable chemicals used in analytical procedures or instruments. There is special concern with toxic volatile waste chemicals that could get into the cabin atmosphere.

4. The SS will have limited amounts of power and space available for any operation, so these factors as well as weight will influence the choice of analytical instruments. Modularity and interchangeability of components between different instruments should be a goal in order to reduce the inventory of spare parts and improve serviceability.

#### DIRECT METHODS

Direct methods are able to continuously monitor the PW stream without personnel involvement. They have a specific response to the parameter of interest and should not respond to other substance. The methods we will consider are conductivity, photometry, and electrochemistry. These methods are most useful for monitoring general or surrogate parameters relative to the function of the PW purification system, though some are applicable to specific substances.

##### Conductivity

This is the basic test for total ion concentration in the water and increasing conductivity indicates the ion exchange resins need renewing, or there has been a malfunction and excess ions are entering the PW system from some source.

Usual equipment: Two platinum electrodes (with Pt black coating) with a fixed area and spacing, energized with a constant small amplitude sine wave voltage and read with a Wheatstone bridge circuit. The output is either in conductivity (units  $\mu\text{S}/\text{cm}$ , previously  $\mu\text{mhos}/\text{cm}$ ) or its reciprocal, as resistivity (ohms cm or megohm cm).

Suggested improvements for SS operation: The commercial equipment can be used "as is" with the only interference coming from gas bubbles in water. However, the system can be improved by using a bipolar pulse detector circuit (4) which allows simple rugged electrodes of materials like stainless steel to be built into the water line (example, ring electrodes separated with Teflon insulating rings). This method is very sensitive and responds rapidly to conductivity changes. It is used in many modern instruments, often with micro cells. The electronics are simple and interface readily to computer control. In principle, one bipolar pulse circuit could be multiplexed to all the conductivity detectors and other resistance devices (like thermistors) on-board. If several units are used they would have the advantage of interchangeability.

##### Photometry

This is a group of methods including turbidimetry, spectrophotometry, and fiber optic sensors (optrodes), all based upon sensing the amount of light that is absorbed or scattered at one or more wavelengths.

Turbidimetry is a required water quality measurement related to the amount of light scattered from particulate matter suspended in water. If the scattered light is actually measured the method is called

nephelometry and is more sensitive for low light scattering. On the SS, increasing turbidity will generally indicate the presence of bacterial growth in the system or breakdown of the activated carbon or ion exchange resins. Detergent molecules, which form micelles, and gas bubbles will also scatter light. This permits the detection of unallowed gas bubbles in the PW as their size, and thus turbidity, is pressure sensitive.

Spectrophotometry is required for the PW "color" requirement. It is also being considered as a surrogate measurement for total organic carbon (TOC) in PW. However, individual organic compounds vary considerably in the amount of light absorbed at a specific wavelength, with most aliphatic compounds only weakly absorptive ( $\epsilon < 40$ ) above 200 nm. The method, if carried out at several specific wavelengths, may give a reasonable measure of aromatic compounds, as they strongly absorb ( $\epsilon > 8000$ ) in the region of 200 to 350 nm. If one makes the assumption that aromatic containing chemicals are the most hazardous, this measurement would be closer to a surrogate response for the "TOC less nontoxicants".

Usual equipment: Photometric processes require either laser or broad band, highly stable, light sources (with monochrometers or filters to select the desired wavelengths) and an optical train free from unwanted stray light and vibrations. The detectors need to be very sensitive with a wide dynamic range of response. This means they need very well regulated power supplies, often with either high voltage or current output capability.

Suggested improvements for SS operation: If several optical measurements are to be made, it may be possible to combine them, using only one or two light sources (UV and Visible) and one monochromator, with fiber optics used to pick up specific wavelengths and distribute them to the sample cells. Also, absorption measurements for contaminants of PW requires 10 cm or longer path length cells. For a typical aromatic compound at 100 ppb, with a molar absorptivity of  $8 \times 10^4$ , a 10 cm cell would give only an absorbance of 0.01 units (only 2.3% of the light absorbed). Other problems for photometric measurements on the SS come from absorbance due to added biocide (forms of  $I_2/I^-$ ), this may need prior removal (see section on "Silver Bullet"), and need to correct for scattering due to gas bubbles or turbidity. Finally to circumvent these problems the use of "optrodes", optical sensors equivalent to electrodes, may be used for some specific determinations. Optrodes have a reagent immobilized on the end of a double fiber optic that gives a color or fluorescence with the species of interest. Light is sent down one fiber and returned up the other and does not enter the solution (5).

### Electrochemistry

Direct measurements with electrochemistry include potentiometry, using ion and molecular selective electrodes (ISE and MSE) and oxidation reduction potential (ORP) electrodes, and some applications of controlled potential amperometry (CPA).

An ISE for pH determination is required for PW, and the activity of several other specific ions could be measured, examples  $I^-$  and  $F^-$ , however many other ISE suffer interference problems. MSEs could be used to measure  $O_2$ , and perhaps  $CO_2$  and  $NH_3$ . They also have the possibility for determining biochemical molecules through immobilized enzymes (6). The problem with most of these ISE and MSE is their need for frequent recalibrations. While not required, an ORP electrode (made of Pt or Au) would be useful for monitoring the redox potential of PW, which would indicate the level of  $I_2$  and/or  $O_2$ . A more negative trend in the potential may indicate anaerobic bacterial growth in the system.

The use of CPA with large area inert electrodes could be valuable for monitoring the concentrations of several classes of organic compounds present in PW, Table 2. In aqueous solutions the range of potentials available for CPA, before  $H_2$  or  $O_2$  formation occurs, is about +1.0 to -1.0 volts. The method might be used as a surrogate method for phenols (MCL of 1 ppb) and halogenated hydrocarbons (MCL of 10 ppb). With porous electrodes CPA could be used to remove the interfering  $I_2/I^-$  from PW before the trace organics are measured and to regenerate  $I_2$  from  $I^-$ .

TABLE 2.- POTENTIALLY DETECTABLE ORGANIC FUNCTIONAL GROUPS BY CONTROLLED POTENTIAL AMPEROMETRY (7).

Oxidizable Groups	Potential (V vs SCE)	Reducible Groups	Potential (V vs SCE)
Azines	> +1.2	Halogens	+0.2 > -0.2
Hydrocarbons	> +1.0	Nitro compds	-0.2 > -0.5
Amides	> +0.5	Diazo compds	-0.2 > -0.6
Amines	+1.0 > +0.5	Ethers	-0.6 > -1.4
Quinolines	+0.6 > +0.2	Esters	-0.8 >
Phenols	+0.5 > 0.0	Aldehydes	-1.1 >
		Ketones	-1.4 >
		Olefins	-1.7 >

(Note: This list represents general trends, as it is by no means complete nor totally accurate based upon my experience. RDG)

Usual equipment: For potentiometric measurements an indicator electrode and a reference electrode are needed, one reference electrode would work for several indicator electrodes if they are in close proximity. A very high impedance potentiometer, equivalent to a pH meter, is needed to sense the voltage without disturbing the measurement. For controlled potential amperometry usually separate reference, counter and working electrodes are needed, the latter two would be made of either Pt or Au. A three electrode potentiostat is usually needed to control and monitor the potential and current.

Suggested improvements for SS operation: Commercial equipment can be used, one precaution is the need for gel filled reference electrodes to avoid bubble formation between the electrode and the salt bridge

junction. All the indicator electrodes could be multiplexed thru one output console if each electrode has its own very high impedance operational amplifier. The problems to overcome in potentiometry are the stability of the electrodes and a means of cleaning and recalibrate them during a mission. If the electrodes are used in-line with the PW, a bypass loop to one or more special calibration solutions might be required. The CPA instrumentation is simple and should be interface thru a computer for digital control of potentials and currents. The electrodes may be subject to long term fouling if bacteria are present and may need periodic cleaning and exchange between missions.

### INDIRECT METHODS

Indirect methods usually involve periodic sampling with chemical treatment, or sample separation followed by determination of the species of interest. The reaction processes to be discussed are TOC and two electrochemical techniques, anodic stripping voltammetry and micro-coulometric titrimetry (an insitu titration method). Also discussed are a semi-micro, zero G reaction system and a method for removing the interference of  $I_2/I^-$ . The separations methods of gas chromatography (GC), supercritical fluid chromatography (SFC), high performance liquid chromatography (HPLC), and CZE will be discussed briefly along with their detectors and required interfaces suitable for microgravity operation.

#### Chemical reactions

The TOC method oxidizes the organic carbon present in a PW sample and measures the  $CO_2$  released. This is an important surrogate measure of the organic quality of PW and should be carried out routinely on the SS.

Usual equipment: Organic matter in water samples (<10 ml) is oxidized in a reactor either by photochemical or thermal decomposition of persulfate, or by some other oxidation catalyst. The  $CO_2$  produced is measured by conductivity or IR, usually after trapping the  $CO_2$  in molecular sieves and purging it to the detector. The methods have problems with incomplete oxidation of refractory organic molecules. High temperature persulfate oxidation appears to give superior results and will be the method discussed. Inorganic carbon,  $HCO_3^-$ , interferes and is removed by prior  $N_2$  purging of the acidified sample. This causes some volatile organics to be removed and they must be trapped and oxidized to  $CO_2$ . These separate measurements of  $CO_2$  are called TIC and TVOC respectively.

Suggested Improvements for SS operation: Since  $CO_2$  is sparged from the reactor with  $N_2$  or air, the commercial method needs adaptation to work in microgravity. The method also consumes sodium persulfate and phosphoric acid solutions that should be considered hazardous in the SS environment. An "orbital reactor" (see indirect electrochemical methods) could be used to handle the gas-liquid separation problem and to conduct the persulfate high temperature oxidation. The acid and persulfate could be generated by standard electrochemically reactions within the reactor to avoid the need for consumables other than a sodium sulfate solution.

Microcoulometric titrimetry of the  $\text{CO}_2$  collected by purge and trap methods would allow for smaller water samples, down to 0.1 ml, than the current IR detector system. The reaction could be carried out in a 0.5 ml cell that would require less heat and smaller volume of consumables. These changes effectively convert the TOC determination to an electrochemical method.

**Electrochemical techniques** often require the removal of interferences due to  $\text{I}_2/\text{I}^-$  and  $\text{O}_2$  at the concentrations expected in PW, as they give appreciable reduction currents at small electrode potentials. I would suggest a pair of porous silver/silver iodide electrodes separated with a Nafion ion exchange membrane that can reduce iodide to very low levels, base upon the applied potential and the AgI solubility product of  $1.5 \times 10^{-16}$  at 25 °C. This "silver bullet" (SB) works on the principle that as the required water passes thru one electrode,  $\text{I}_2$  will react spontaneously with Ag to form AgI and  $\text{I}^-$  and further anodic oxidation of Ag will form more AgI on the electrode. Meanwhile, as a small amount of water passes thru the other electrode, its coating of AgI is reduced to Ag and  $\text{I}^-$  is released (this concentrated iodide solution can be collected for re-use in the PW loop after oxidation to  $\text{I}_2/\text{I}^-$ ). The relative amount of  $\text{I}^-$  in the two water streams is controlled by the applied voltage, with 350 mV giving a ratio of  $10^6$ .

Note, the SB method can be readily scaled up to remove  $\text{I}^-$  from PW for end uses. PW at 1 liter per minute would need a maximum current of 190 mA (27 mA nominal, 3 ppm  $\text{I}_{\text{total}}$ ). In this use the SB would have three connection positions, then as the AgI builds up on one electrode and is depleted on the other they can be interchanged so the opposite water streams flow thru the electrodes and their functions are exchanged. The third position would allow bypassing the electrodes to periodically disinfect the terminal portion of the water system with biocide.

Removal of  $\text{O}_2$  from water can also be accomplished by electrochemical reduction in similar flow cells with Pt electrodes, or by sparging the solution with pure nitrogen or argon. An "orbital reactor" (OR) is suggested as a microgravity method for the mixing and sparging of solutions required for chemical or electrochemical reactions. The OR is a platform holding a reaction vessel that is given a orbital motion while its orientation remains fixed, the only moving parts associated with the reactor are the connecting wire leads and tubing. These need to flex with the motion of the OR and their position and composition must be selected to reduce stress fatigue. The reaction cells can be of numerous sizes and shapes, but for electrochemistry a disk or washer shaped cell of 0.5 ml to 5 ml should work well. For example, a washer shaped cell 0.25 mm thick with radii of 1 and 2 cm, the volume would be 0.5 ml and with an orbiting rate of 212 rpm the force would be 0.5 and 1 G at the inner and outer surfaces respectively. For purging the OR cell, gas would be introduced at the periphery of the cell and it would move to the center in a spiral motion. The shear of moving water by the gas inlet would cause the bubbles to be small and well separated, giving very efficient sparging. The optimum amount of G force needs to be

determined, as little as 0.1 to 0.2 G may be sufficient for these operations. The flat shape and large area of these OR cells allows the placement of a variety of sensors in the same cell, for example, conductivity, large area, micro, reference, counter, and thin layer electrodes of several different materials, as well as optrodes and thermistors. Divided electrochemical cells for microcoulometric titrations, where the reagent is generated in situ, can be made by splitting the OR reactor horizontally with an ion exchange membrane.

Anodic stripping voltammetry is a common method for measuring trace (to sub ppb) levels of the heavy metal ions including Cd, Cu, Pb, Hg, Ni, Ag, Fe, Zn, and possibly Cr, Mn, and As in water. This method would replace most of the water analyses done by ground based graphite furnace atomic absorption. This method concentrates the metals by plating them out on an inert electrode, held at a sufficiently negative potential, from a uniformly stirred solution for a specific period of time. After this time the stirring is stopped and when the solution comes to rest the metals are removed by a positive voltammetric sweep. The amount of charge required to remove each metal at its specific redox potential corresponds to the metals concentration in solution. Often the metals are plated into a co-deposited mercury film and the procedure may require the addition of supporting electrolyte.

Microcoulometric titrimetry of the acid and base content of PW can be carried out in a divided OR cell which allows the production of either  $H^+$  or  $OH^-$ , in the PW sample, depending upon the electrode polarity. By applying constant current pulses of short duration, specific amounts of acid or base are generated, and with the use of a sensitive bipolar pulse conductivity detector, the total acid and base content of the PW can be determined from titration curves. The detection limits should, in theory, be at the sub ppb level for acid and base compounds, as a  $1 \mu A$  pulse for 100 msec is equivalent to 0.1 microcoulomb or 0.2 ppb for a 100 molecular weight acid or base in 0.5 ml. This technique might be developed to give a surrogate measure of organic acids, bases and possibly phenols in PW. It might also be used for measuring  $CO_2$  and  $NH_3$  in the PW and overcoming their analytical interference. As mentioned earlier the technique could be adapted to do the TOC analysis. The method could also be used in other SS chemical analyses where a needed titrant can be generated microcoulometrically.

### Separation Techniques

There are four general types of high resolution separation techniques that may be considered for identifying and quantitating components (mainly organic) of PW and other samples on the SS. Three are based upon pressurized fluid elution: GC, generally using nitrogen gas, SFC, using  $CO_2$  above its critical pressure ( $>1100$  psi), and several variations of HPLC, using aqueous/organic solvent mixtures or aqueous buffers (to pressures of 6000 psi). The other one, CZE, uses very high electric field driven elution and comes in several variations. For analyzing trace organics in PW these techniques will probably require concentration

of samples using some arrangement of purge and trap.

The GC technique is excellent for separating volatile heat stable organic compounds and interfaces well to a variety of detectors. It is considered the main tool for cabin atmosphere monitoring and could be useful for some specific PW analyses.

Usual equipment: The method requires a source of high pressure gas (100 psi) a sample injection system (sampling loop and valve), a temperature programmable oven (to 350 °C), a long silica capillary column with a bonded stationary phase (several may be needed) and an interface to a detector.

Suggested improvements for SS operation: Most commercial equipment would be compatible with the limitations of the SS environment. The small N<sub>2</sub> requirement can be met by using cabin make-up gas if properly purified, or generated electrochemically from hydrazine sulfate (toxic). The standard flame ionization detector (FID) should be readily interfaced to the column if a premix H<sub>2</sub>/O<sub>2</sub> micro burner is used to avoid microgravity effects. Otherwise a phototization detector (PID) might be adequate. A fourier transform infrared detector (FTIR) should work directly, but the electrolytic conductivity detector (ELCD) that is excellent for halogen compounds would need modification of its gas/liquid reactor. The mass spectra detector (MSD), which will be discussed later, would interface easily.

The SFC technique is still developing and currently is able to separate heat sensitive compounds which are non-volatile, if they are not too polar, as well as the compounds handled by capillary GC. The technique falls between GC and HPLC in its applications and interfaces to most of the detectors used for GC.

Usual equipment: The equipment is usually a modification to either GC or HPLC instruments, and requires high pressure pumps (often the syringe type), usually a column oven for temperature up to 150 °C and an interface to a detector.

Suggested improvements for SS operation: Of the three chromatographic techniques this is the most compatible with SS operation since it uses CO<sub>2</sub> (which is recycled on the SS and is non-hazardous) and has modest power requirements. An advantage of supercritical CO<sub>2</sub> is its ability to purge traps of all but their most polar organic compounds. The eluting ability of CO<sub>2</sub> can be modified by changing its density (pressure) and by saturating it with water or polar organics. SFC has about the same detector limitations as GC.

The HPLC techniques have reached a high degree of sophistication and the ones of interest are normal and reverse phase partitioning for organics (NPP & RPP), ion pairing for ionic and ionizable organics (IP), and ion exchange, mainly for inorganic ions, called ion chromatography (IC). In combination these techniques have the ability to separate most mixtures

of organic compounds including the most polar and non-volatile, though the resolution is not as high as with GC or SFC.

**Usual equipment:** The methods all require special very high pressure pumps and controllers to push eluent solutions rapidly thru short (10 cm) small bore columns packed with special very small diameter particles designed or coated for the particular application. This means that either dedicated instruments are needed for each application or columns and eluent systems must be changed for different analyses. For each sample the NPP and RPP systems use about 10 to 15 ml of a complex eluent mixture, often containing acetonitrile or methanol as a major component. The IC requires several aqueous ionic buffer combinations and ion exchange columns to separate the different sets of ions effectively with a sub ppb level of sensitivity. The IP technique is a combination of the previous two and requires an organic solvent with ionic buffer and separation of organic ion pairs is usually on RPP columns. The detectors for these methods are generally ones that operate well with solutions, for example, conductivity, electrochemical (often CPA), and various photometric methods. Recently, by means of a complex interfaces, several types of mass spectra detectors (MSD) have been successfully used with HPLC (see the section on detectors).

**Suggested improvements for SS operation:** The large amount of consumable buffers and their waste, some of which are toxic, may limit the use of HPLC in the SS. The amount of eluent can be reduced about 80 fold by using "microbore" columns (8). However, the IC which is the most important for PW analysis apparently does not have the extra small particle ion exchange resins needed for microbore modification at this time.

**The CZE techniques** are very new and hold much promise as a universal separation technique (9). They use the principle of electroosmotic force to move buffer solutions down small bore (<100  $\mu$ m) while at the same time using the electrophoretic mobility to separate charged species. Besides zone electrophoresis itself, there are two variant techniques, isotacophoresis that uses a discontinuous ionic buffer to focus and separate analyte ions, and electrokinetic chromatography that uses an ionic micelle buffer to separate neutral organic molecules.

**Usual equipment:** No commercial equipment is currently available, but the apparatus is simple to build. The methods requires a 30 kV DC at 10  $\mu$ A power supply, a 0.5 to 2 m length of fused silica capillary tubing, a very small volume of aqueous buffers, and an interface to a detector. Sample size for CZE is about 10 nl, and the resolution is about equivalent to capillary GC. Because of the small sample size and column diameter it requires very sensitive and high resolution detectors.

**Suggested improvements for SS operation:** The techniques are very compatible with the SS environment, with small weight, volume, power and consumables (non-hazardous) requirements. However, considerable development is needed to find the right buffer protocols to accomplish

the required separations. For convenience of SS operation, several columns and buffer setups may be required that can be switched between one power supply. It may be necessary to change capillary columns frequently due to plugging or other failure and replacements with easy means of installation is needed. Samples of PW will have to be concentrated for use with these techniques. For the required sub  $\mu$ l sized samples, this might be accomplished without loss of too many volatile organics by diffusion of water out thru a membrane.

## Detectors

A number of detectors are available for the separation methods just discussed, and some were commented upon in that section. The mass spectra detector (MSD) in its several variations is close to the universal detector, giving nearly absolute identification of unknown components, requiring very small samples, having very high sensitivity, and interfacing easily to GC and SFC. The interface to HPLC and CZE is more difficult because the eluent is a liquid. Several techniques have been developed, but there are problems with getting all analyte species ionized and into the MSD with equal efficiency. Electrochemical and conductivity detectors interface readily to the HPLC and CZE columns and are very sensitive and can be made with very small volume cells. There are several other detectors which may be useful but will not be discussed: photoionization, UV absorption, fluorescence and infrared.

The MSD methods include the low resolution units (Ion Trap, Mass Sensitive Detector and related ion drift methods) that give only the parent mass/charge ion of unit resolution, and the high resolution mass spectra systems (quadrupole units are the smallest and simplest) that give a complete molecular formula and often an unambiguous structure. Except for the ion drift method, the detectors all require a very good quality vacuum ( $10^{-7}$  Torr) and need high velocity pumping for interfacing to HPLC or CZE. The requirements of complexity, weight, volume, and energy consumption make the MSD the least compatible of the analytical techniques for SS, however their advantage may be worth the effort to develop a workable system. The most likely candidates would be an ion drift, or ion trap detector. The ion trap is the smallest detector, excluding the turbomolecular pump, it is about 64 cubic inches, while the ion drift unit works with a plasma at atmospheric pressure.

The MSD interface requires removing the volatile part of the eluent stream from the analyte species and converting them to "dry" ions while introducing them into the mass/charge detector system. Two interfaces currently in use are electrospray that uses high voltage to nebulize and ionize the droplets and thermospray that uses rapid heating of the effluent containing special buffers (usually  $\text{NH}_4\text{HCO}_3$ ) to produce ionization under these conditions. The removal of solvent is aided by a rapid flow of dry  $\text{N}_2$  in front of the MSD inlet. There is still a need for research on these interface methods and a combination of the two may be needed for the special conditions on the SS.

Electrochemical and conductivity detectors can both use the same microcell and electrode design to interface efficiently to HPLC and CZE systems. The conductivity detector would use the same bipolar pulse circuit that has been indicated for other conductivity measurements. The electrochemical detector would normally operate as a controlled potential amperometer (CPA), but could be used for rapid scan cyclic voltammetry for unknown identification. The micro volume cell that would fit directly on the end of either microbore HPLC or CZE columns could be constructed in sandwich fashion from thin Pt or Au foils (0.01 mm) between thin Teflon films (0.03 mm). A hole to match the inside diameter of the column (0.1 mm or less) would be punched or laser drilled through the layers to make a set of thin ring electrodes. The volume of the cell would be about 0.3 nl and would not degrade the resolution of a CZE column. Except for identifying unknowns, where accurate mass and structural information is needed, these detectors would work very well for most HPLC or CZE separation on the SS.

#### RECOMMENDATIONS

A variety of suggestions have been made throughout this report to improve the analytical techniques for SS operation. The most important recommendations are:

- (1) The silver/silver chloride electrode (SB) method of removing  $I_2/I^-$  biocide from the water, since it may interfere with analytical procedures for PW and also its end uses (page 13-10).
- (2) The orbital reactor (OR) method of carrying out chemistry and electrochemistry in microgravity by using a disk shaped reactor on an orbital table to impart artificial G force to the contents, allowing solution mixing and separation of gases and liquids (page 13-10).
- (3) A simple ultra low volume highly sensitive electrochemical/-conductivity detector for use with a capillary zone electrophoresis apparatus (page 13-15).

It is also recommended, since several different conductivity and resistance measurements are made during the analysis of PW, that the bipolar pulse measuring circuit be used in all these applications for maximum compatibility and redundancy of equipment (page 13-5).

#### REFERENCES

1. Space Station Program Definition and Requirements Document. NASA SST 30000. Rev D., 1987, Table 2-34.
2. Shrader Analytical and Consulting Laboratories Report of Analytical Services: Space Lab 3 ARS Humidity Condensate. SL 12352, 1985.
3. Sorbert Purification Testing of Air Revitalization Condensate. JSC 19170-SS-089 (CSD), 1985.
4. Johnson, D. E. and C. G. Enke: Bipolar Pulse Technique for Fast Conductance Measurements. Analytical Chemistry, vol. 42, March 1970, pp. 329-335.
5. Seitz, W. R.: Chemical Sensors Based on Fiber Optics. Analytical Chemistry, vol. 56, January 1984, p. 16A.
6. Rechnitz, G.: Bioanalysis with Potentiometric Membrane Electrodes. Analytical Chemistry, vol. 54, September 1982, p. 1194A.
7. Skoog, D. A.: Principles of Instrumental Analysis, third edition. Saunder College Publishing. 1985, p. 797.
8. Isco Application Lab Staff: Enhanced Mass Sensitivity and Resolution in Microbore LC by Improvements in Detection and Liquid Flow Path. Isco Applications Bulletin, No. 44, March 1985, p.4.
9. Smith, R. D.; Olivares, J. A.; Nguyen, N. T.; Udseth, H. R.: Capillary Zone Electrophoresis-Mass Spectrometry Using an Electrospray Ionization Interface. Analytical Chemistry, vol. 60, March 1988, pp.436-441.

MODEL EVALUATION, RECOMMENDATION, AND  
PRIORITIZING OF FUTURE WORK FOR THE  
MANIPULATOR EMULATOR TESTBED

Final Report

NASA/ASEE Summer Faculty Fellowship Program--1988

Johnson Space Center

Prepared By:	Frederick A. Kelly, Ph.D.
Academic Rank:	Assistant Professor
University & Department:	Michigan Technological University Mechanical Engineering - Engineering Mechanics Department Houghton, Michigan 49931
NASA/JSC	
Directorate:	Engineering
Division:	Systems Development and Simulation
Branch:	Teleoperator Systems
JSC Colleague:	Edith C. Taylor
Date Submitted:	August 12, 1988
Contract Number:	NGT 44-005-803

## ABSTRACT

The Manipulator Emulator Testbed (MET) is to provide a facility capable of hosting the simulation of various manipulator configurations to support concept studies, evaluation, and other engineering development activities. Specifically, the testbed is intended to support development of the Space Station Remote Manipulator System (SSRMS) and related systems. The MET is required to permit that components simulated in software may be replaced by future hardware components.

The objective of this study is to evaluate the math models developed for the MET simulation of a manipulator's rigid body dynamics and the servo systems for each of the driven manipulator joints. Specifically, the math models are examined with regard to their amenability to pipeline and parallel processing processing. Based on this evaluation and the project objectives, a set of prioritized recommendations are offered for future work.

## INTRODUCTION

The Manipulator Emulator Testbed (MET) is to provide a facility in which different manipulator configurations may be simulated. The MET will be used as a tool to support Space Station manipulator development. It will be used to develop and evaluate concepts, support design and development studies, and evaluate hardware components and flight software modules. The MET is required to be designed such that initial software simulated components may be replaced by future hardware elements [1] - [3].

The MET is currently built around a network of IBM PC-AT computers. Each computer operates at an 8-MHz clock rate and is equipped with an Intel 80287 math co-processor, 640K bytes of memory, an extended graphics adapter (EGA) card, a color monitor, a 30-MB hard disk drive, and a high-capacity floppy disk drive. Each PC is equipped with a National Instruments GPIB-PCAA IEEE-488 interface card. The network includes an operator control station, data recording and display capability, and hardcopy output. A Multibus II "Network-in-a-Box" is scheduled to be added to the MET in August 1988.

The software includes the operating system (iRMK), intercomputer communication software from National Instruments, external interface, executives, operator support, data recording, and math model modules. The math model application software includes (1) a multi-link manipulator rigid body dynamics model and (2) joint servo models for each of the driven manipulator joints. The MET software is intended to provide a simulation of proposed manipulator designs. The math model software is distributed across the PC's such that a hardware component may be substituted and consequently the math model removed.

The objective of this study is to evaluate the math models developed for the MET simulation of a manipulator's rigid body dynamics and the servo systems for each of the driven manipulator joints. Specifically, the math models were examined with regard to their amenability to pipeline and parallel processing. Based on this evaluation and the project objectives, a set of prioritized recommendations are offered for future work.

## MODEL EVALUATION

A recursive rigid body dynamics formulation was developed for real-time simulation on the MET by Nasser [4]. Nasser considers it to be a generalization of a method due to Armstrong [5]. The formulation is general in the sense that translational or rotational joint types can be modeled. Nasser also claims that the model assumes a topological tree configuration, but doesn't describe how this would be done. Procedures described in [6] - [7] are particularly suitable to extending Nasser's method to general topological trees.

The procedure for solving for the reaction force and torques at each joint and the joint accelerations given the external forces and moments on each link and the actuator forces and moments can be described by referring to Figure 1 [4].

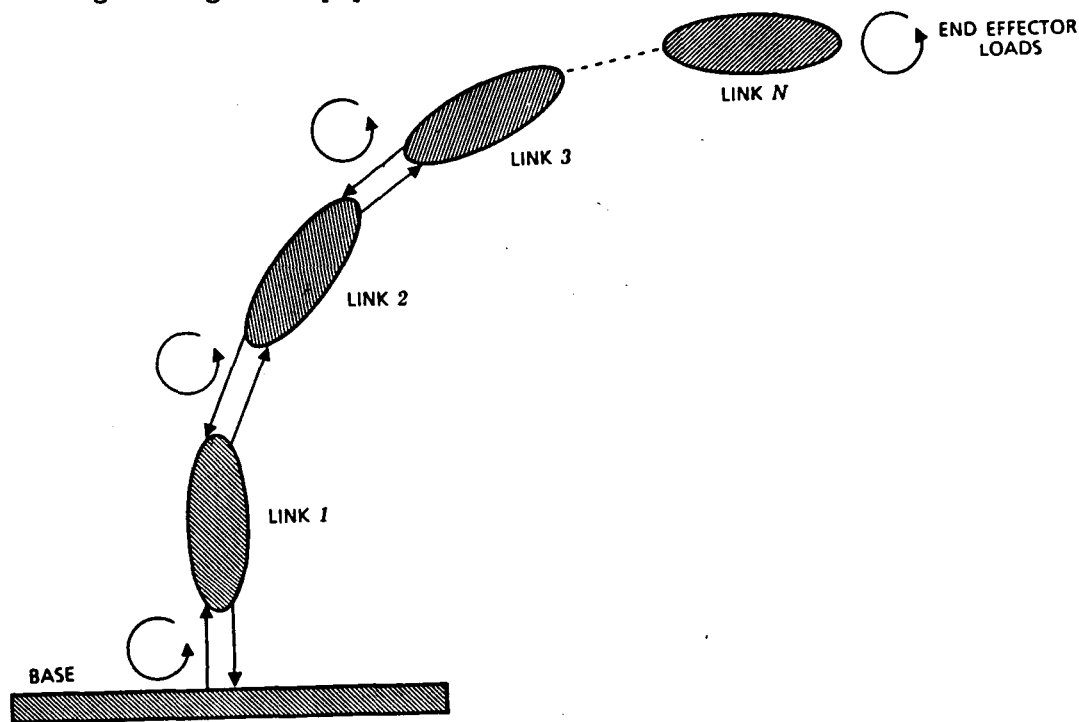


Figure 1. Open Kinematic Chain

First, set up the equations of dynamic equilibrium for link N and then solve for the reaction forces and moments and the joint acceleration in terms of the remaining variables. The remaining variables include the proximal links' joint displacements velocities and accelerations and the distal links' applied forces and moments, displacements, and velocities. Next, proceed to link N-1 and substitute the expressions for the reactions that involve link N-1 joint accelerations into the equations of dynamic equilibrium for link N-1. The reaction forces and moments exerted by link N-2 on link N-1 and the joint acceleration of link N-1 should then be solved for in terms of the remaining variables. This step is repeated until link 1 is reached. The joint acceleration for link 1 is then obtained in terms of the distal links' external forces and moments, actuator forces and moments and state variables, the state of link 1, and the base state. Then move distally to link 2 and substitute for the link 1 joint acceleration to find the joint 2 acceleration. This step is repeated until link N is reached and all the joint accelerations are found.

A free body diagram for any link i of the open kinematic chain is shown in Figure 2.

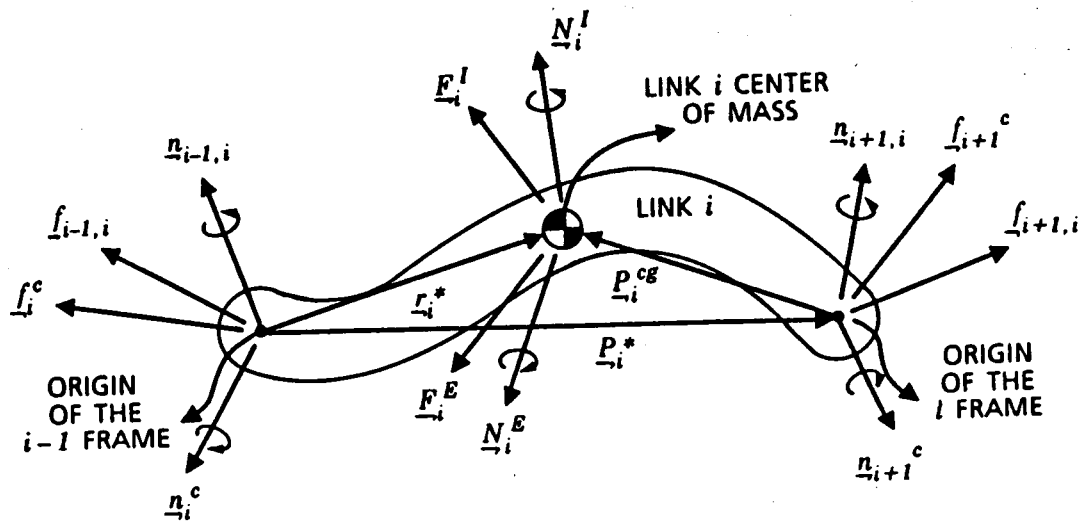


Figure 2. Link Free Body Diagram

where

- $\underline{F}_i^I$  is the link i inertia force,
- $\underline{N}_i^I$  is the link i inertia torque,
- $\underline{f}_{i-1}^c$  is the control force exerted on link i at joint i-1,
- $\underline{n}_{i-1}^c$  is the control torque exerted on link i at joint i-1,
- $\underline{f}_{i+1}^c$  is the control force exerted on link i+1 at joint i,
- $\underline{n}_{i+1}^c$  is the control torque exerted on link i+1 at joint i,
- $\underline{f}_{i+1,i}$  is the reaction force that link i+1 exerts on link i,
- $\underline{n}_{i+1,i}$  is the reaction moment that link i-1 exerts on link i
- $\underline{f}_{i-1,i}$  is the reaction force that link i-1 exerts on link i,
- $\underline{n}_{i-1,i}$  is the reaction torque that link i-1 exerts on link i,
- $\underline{F}_i^E$  is the sum of the external forces acting on link i referred to the mass center of link i,
- $\underline{N}_i^E$  is the sum of the external moments acting on link i when the external forces are referred to the mass center of link i,
- $\underline{r}_i^*$  is the position vector of the mass center of link i relative to the origin of the i-1 frame,

$P_{-1}^{CG}$  is the position vector of the mass center of link  $i$  relative to the origin of the  $i$  frame,

$P_{-1}^*$  is the position vector of the origin of frame  $i$  relative to the origin of frame  $i-1$

A number of researchers have investigated pipeline and parallel implementations of the equations of motion for various purposes. Luh and Lin [8] described a procedure for scheduling the subtasks of a group of microprocessors computing the inverse dynamics using the Newton-Euler equations of motion. One microprocessor is assigned to each manipulator link. A variable branch-and-bound search finds an optional subtask-ordered schedule for each of the microprocessors. However, the total processing time of solving the minimum-time scheduling problem could not be easily reduced to a manageable level.

Lathrop [9] proposed two parallel algorithms for solving the inverse dynamics problem using a group of special-purpose processors. One is a linear Newton-Euler algorithm, which is most closely related to the method proposed by Luh and Lin. The other is a logarithmic parallel Newton-Euler algorithm based on the "partial sum" technique, which achieves on  $O(\lceil \log_2 N \rceil)$  total processing time. However, both methods have two main effects that deteriorate the performance of parallel computations. They both require potentially massive internal buffering to achieve pipelined computation between forward and backward recursions. They both involve complex intercommunication and bussing, which frequently cause data to be fetched, and as a result data for operand pairs are not properly aligned for parallel computations.

Lee and Chang [10] proposed an algorithm for solving the inverse dynamics problem of an  $N$ -link manipulator using  $p$  processors in parallel on the Newton-Euler equations of motion. It was the "recursive doubling" technique to achieve a total processing time of  $O(k_1 \lceil N/p \rceil + k_2 \lceil \log_2 p \rceil)$ , where  $k_1$  and  $k_2$  are constants. When  $p = N$ , the algorithm is  $O(\lceil \log_2 N \rceil)$ , the same as Lathrop. The  $p$ -fold parallel algorithm consists of  $p$ -parallel blocks with pipelined elements within each parallel block. The results from the  $p$ -parallel blocks form a homogeneous linear recurrence of size  $p$ . The parallel algorithm can also be implemented in a systolic pipelined architecture, requiring three floating-point operations for each complete set of joint torques.

Binder and Herzog [11] described an algorithm based on the recursive Newton-Euler equations of motion where the parallel computations are distributed over multiple computing elements, one for each joint. Concurrency is achieved by substituting "predicted" values for the actual values of variables involved in the recursive equations. The authors simulated this method and compared it to other approaches such as simplification of the dynamic equations.

Amin-Javaheri and Orin [12] proposed systolic architectures consisting of 1,  $N$ , and  $N(N+1)/2$  processors for computing the inertia matrix of an  $N$  degree of freedom manipulator. A VLSI-based Robotics Processor is being developed as the fundamental component of the

architecture. The algorithm used is based on recursive computation of the inertial parameters of sets of composite rigid bodies and is programmed to exploit any inherent parallelism.

Lee and Chang [13] proposed two parallel algorithms for computing the forward dynamics for real-time simulation with N processors for an N degree of freedom manipulator. The first algorithm is based on Walker and Orin's [14] composite rigid-body method. It generates the inertia matrix using the parallel Newton-Euler algorithm, the parallel linear recurrence algorithm [10], and the modified row-sweep algorithm, and then inverts the inertia matrix to obtain the joint acceleration vector. The time complexity of this parallel algorithm is of the order  $O(N^2)$  with  $O(N)$  processors. Further reduction of the order of time complexity can be achieved by implementing the Cholesky factorization procedure on VLSI array processors to invert the symmetric, positive-definite, inertia matrix. The second parallel algorithm is based on Walker and Orin's [14] conjugate gradient method.

Of the methods just described, only Nasser's method [4] was formulated to allow for hardware substitution in a simulation. Currently, it is implemented on the MET in a sequential algorithm with one IBM PC-AT. The Nasser method appears to be a variation of a method by Featherstone [15]. The  $A_i^*$  matrix in Nasser's method seems to correspond to the  $\hat{I}_i^A$  matrix in Featherstone's method. This would imply that Nasser's method has the same nonlinear recurrence that makes Featherstone's method impossible to solve with known methods of parallel solution of recurrence problems [13]. Kung [16] has shown that any parallel algorithm using any number of processors cannot be essentially faster than the obvious sequential algorithm for any nonlinear polynomial recurrence problem. If not for this, the recursive doubling technique of Kogge and Stone [17] could have been used. The method of Nasser does not seem to be particularly adaptable to pipelining either. The calculation of the  $A_i^*$  matrix alone requires about 42% of the total computations. The other processors in the pipeline would have to be idle, waiting for the  $A_i^*$  to complete.

The hardware that makes up the MET limits the use of pipelining. Instruction pipelining is possible on the Intel 80286 only by prefetching instructions. As an alternative, the MET would seem to be an ideal application for the new Intel 80960 architecture which allows parallel and out-of-order instruction execution of 32-bit operations at 20 MHz [18].

The Multibus II "Network-in-a-Box" promises to greatly speed up interprocessor communications with initially three Intel 80386 processors on board. The National Instruments IEEE-488 interface card and software are much too slow.

#### RECOMMENDATION AND PRIORITIZING OF FUTURE WORK

My recommendation of future work for the MET will be listed in order of decreasing priority.

1. The Nasser algorithm does not appear to be amenable to pipeline or parallel processing. One option would be to run it as a very fast sequential algorithm on a fast processor. Another option would be to modify the method of Binder and Herzog [11] and use prediction to estimate values, instead of waiting for a complete set of computed values to be ready.

2. The simulation should include the rigid body dynamics of the base link. A mobile base, where motion relative to either a fixed base or base with rigid body dynamics should be included in the model.

3. The servo systems are decoupled in their current design. This makes these computations naturally parallel. The rigid body model of Nasser could run on a fast processor and broadcast results to N processors, one for each of the N servo systems, operating in parallel.

4. Flexible links should be added to the model. For open topological trees, this should be straight forward since the complete force system acting on each link is known. A beam element model of a flexible manipulator is described in Kelly and Huston [19].

5. Constrained motion should be modeled so that closed loops can be simulated. Rigid links should be simulated first and then flexible links.

## REFERENCES

1. Van Valkenburg, J. A.: Manipulator Emulator Testbed Development Project Plan. LEMSCO-23536, January 1987.
2. Schindeler, R. E.: Manipulator Emulator Testbed Verification Plan. LEMSCO-24095, October 1987.
3. Rob, K. K.: Manipulator Emulator Testbed Verification Report. LEMSCO-24113, May 1988.
4. Nasser, M. G.: Recursive Newton-Euler Formulation of Manipulator Dynamics. LEMSCO-24673, February 1988.
5. Armstrong, W. W.: Recursive Solution to the Equations of Motion of an N-Link Manipulator. Proceedings of the 5th World Congress on the Theory of Machines and Mechanisms, Volume 2, Montreal, July 1979, pp. 1343-1346.
6. Huston, R. L., Passerello, C. E., and Harlow, M. W.: Dynamics of Multirigid-Body Systems. Trans. SME, J. Applied Mechanics, Volume 45, December 1978, pp. 889-894.
7. Kelly, F. A.: On the Dynamics of Flexible Multibody Systems. Ph.D. Dissertation, University of Cincinnati, June 1982.
8. Luh, J. Y. S., and Lin, C. S.: Scheduling of Parallel Computation for a Computer-Controlled Mechanical Manipulator. IEEE Trans. on Systems, Man, and Cybernetics, Volume SMC-12, No. 2, March/April 1982.
9. Lathrop, R. H.: Parallelism in Manipulator Dynamics. The International Journal of Robotics Research, Volume 4, No. 2, Summer 1985, pp. 80-102.
10. Lee, C. S. G., and Chang, P. R.: Efficient Parallel Algorithm for Robot Inverse Dynamics Computation. IEEE Trans. on Systems, Man, and Cybernetics, Vol. MC-16, No. 4, July/August 1986, pp. 532-542.
11. Binder, E. E., and Herzog, J. H.: Distributed Computer Architecture and Fast Parallel Algorithms in Real-Time Robot Control. Ibid., pp. 543-549.
12. Amin-Javaheri, M., and Orin, D. E.: A Systolic Architecture for Computation of the Manipulator Inertia Matrix. Proceedings of the 1987 IEEE International Conference on Robotics and Automation, Vol. 2, pp. 647-653.
13. Lee, C. S. G., and Chang, P. R.: Efficient Parallel Algorithms for Robot Forward Dynamics Computation. IEEE Trans. on Systems, Man, and Cybernetics, Volume 18, No. 2, March/April 1988.

14. Walker, M. W., and Orin, D. E.: Efficient Dynamic Computer Simulation of Robot Mechanisms. Trans. ASME J. Dynamic Systems, Measurement and Control, Volume 104, September 1982, pp. 205-211.
15. Featherstone, R.: The Calculation of Robot Dynamics Using Articulated-Body Inertia. The International Journal of Robotics Research, Volume 2, No. 1, Spring 1983, pp. 13-30.
16. Kung, H. T.: New Algorithms and Lower Bounds for the Parallel Evaluation of Certain Rational Expressions and Recurrences. Journal of the Association for Computing Machinery, Volume 23, No. 2, April 1976, pp. 252-261.
17. Kogge, P. M., and Stone, H. S.: A Parallel Algorithm for the Efficient Solution of a General Class of Recurrence Equations. IEEE Trans. on Computers, Volume C-22, No. 8, August 1973, pp. 786-793.
18. Ryan, D. P.: Intel's 80960: An Architecture Optimized for Embedded Control. IEEE Micro, June 1988, pp. 63-76.
19. Kelly, F. A., and Huston, R. L.: Statics and Dynamics of a Flexible Manipulator. Proceedings of the 5th ASME International Computers in Engineering Conference, Boston, MA, 1985.

## REPORT DOCUMENTATION PAGE

1. Report No. <b>NASA CR 172118</b>		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle <b>NASA/ASEE Summer Faculty Fellowship Program--1988 Volume 1</b>				5. Report Date <b>February 1989</b>	
				6. Performing Organization Code	
7. Author(s) <b>Richard B. Bannerot and Stanley H. Goldstein, Editors</b>				8. Performing Organization Report No.	
				10. Work Unit No.	
9. Performing Organization Name and Address <b>The University of Houston--University Park Houston, Texas 77004</b>				11. Contract or Grant No. <b>NGT-44-005-803</b>	
				13. Type of Report and Period Covered <b>Contractor Report</b>	
12. Sponsoring Agency Name and Address <b>National Aeronautics and Space Administration Washington, D.C. 20546</b>				14. Sponsoring Agency Code	
				15. Supplementary Notes	
16. Abstract <p>The 1988 Johnson Space Center (JSC) National Aeronautics and Space Administration (NASA)/American Society for Engineering Education (ASEE) Summer Faculty Fellowship Program was conducted by the University of Houston and JSC. The 10-week program was operated under the auspices of the ASEE. The program at JSC, as well as the programs at other NASA Centers, was funded by the Office of University Affairs, NASA Headquarters, Washington, D.C. The objectives of the program, which began in 1965 at JSC and in 1964 nationally, are (1) to further the professional knowledge of qualified engineering and science faculty members; (2) to stimulate an exchange of ideas between participants and NASA; (3) to enrich and refresh the research and teaching activities of participants' institutions; and (4) to contribute to the research objectives of the NASA Centers.</p> <p>Each faculty fellow spent 10 weeks at JSC engaged in a research project commensurate with his/her interests and background and worked in collaboration with a NASA/JSC colleague. This document is a compilation of the final reports on the research projects done by the faculty fellows during the summer of 1988. Volume 1 contains reports 1 through 14, and volume 2 contains reports 15 through 26.</p>					
17. Key Words (Suggested by Author(s))			18. Distribution Statement <b>Unclassified-Unlimited</b>		
19. Security Classification (of this report) <b>Unclassified</b>		20. Security Classification (of this page) <b>Unclassified</b>		21. No. of pages <b>371</b>	
				22. Price <b>NTIS</b>	